CEC Theses and Dissertations

College of Engineering and Computing

2015

Image Enhancement of Cancerous Tissue in Mammography Images

Richard Thomas Richardson
Nova Southeastern University, richrich@nova.edu

This document is a product of extensive research conducted at the Nova Southeastern University College of Engineering and Computing. For more information on research and degree programs at the NSU College of Engineering and Computing, please click here.

Follow this and additional works at: http://nsuworks.nova.edu/gscis_etd

Part of the Cancer Biology Commons, Computer Sciences Commons, Diagnosis Commons, Health Information Technology Commons, Oncology Commons, Radiology Commons, and the Women's Health Commons

Share Feedback About This Item

Image Enhancement of Cancerous Tissue in Mammography Images

by

Richard Thomas Richardson

A dissertation proposal submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in
Computer Science

Graduate School of Computer and Information Sciences
Nova Southeastern University

2015

This page will be generated by the Program Office

UMI Microform

We hereby certify that this dissertation, submitted by Richard Thomas Richardson, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirements for the degree of Doctor of Philosophy.

_____   _____
Michael Laszlo, Ph.D.   Date
Chairperson of Dissertation Committee

_____   _____
Wei Li, Ph.D.   Date
Dissertation Committee Member

_____   _____
Francisco Mitropoulos, Ph.D.   Date
Dissertation Committee Member

Approved:

_____   _____
Eric Ackerman, Ph.D.   Date
Dean, Graduate School of Computer and Information Sciences

Graduate School of Computer and Information Sciences
Nova Southeastern University
2015

An Abstract of a Dissertation Submitted to Nova Southeastern University

in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Image Enhancement of Cancerous Tissue in Mammography Images

by
Richard Thomas Richardson
December 2015

This research presents a framework for enhancing and analyzing time-sequenced mammographic images for detection of cancerous tissue, specifically designed to assist radiologists and physicians with the detection of breast cancer. By using computer aided diagnosis (CAD) systems as a tool to help in the detection of breast cancer in computed tomography (CT) mammography images, previous CT mammography images will enhance the interpretation of the next series of images. The first stage of this dissertation applies image subtraction to images from the same patient over time. Image types are defined as temporal subtraction, dual-energy subtraction, and Digital Database for Screening Mammography (DDSM). Image enhancement begins by applying image registration and subtraction using Matlab 2012a registration for temporal images and dual-energy subtraction for dual-energy images. DDSM images require no registration or subtraction as they are used for baseline analysis. The image data are from three different sources and all images had been annotated by radiologists for each image type using an image mask to identify *malignant* and *benign*.

The second stage involved the examination of four different thresholding techniques. The *amplitude* thresholding method manipulates objects and backgrounds in such a way that object and background pixels have grey levels grouped into two dominant and different modes. In these cases, it was possible to extract the objects from the background using a threshold that separates the modes. The *local thresholding* introduced posed no restrictions on region shape or size, because it maximized edge features by thresholding local regions separately. The overall *histogram* analysis showed minima and maxima of the image and provided four feature types—mean, variance, skewness, and kurtosis. *K-means* clustering provided sequential splitting, initially performing dynamic splits. These dynamic splits were then further split into smaller, more variant regions until the regions of interest were isolated. *Regional-growing* methods used recursive splitting to partition the image top-down by using the average brightness of a region. Each thresholding method was applied to each of the three image types.

In the final stage, the training set and test set were derived by applying the four thresholding methods on each of the three image types. This was accomplished by running Matlab 2012a grey-level, co-occurrence matrix (GLCM) and utilizing 21 target feature types, which were obtained from the Matlab function texture features. An additional four feature types were obtained from the state of the histogram-based features types. These 25 feature types were applied to each of the two classifications *malignant*

and *benign*. WEKA 3.6.10 was used along with classifier J48 and cross-validation 10 fold to find the precision, recall, and f-measure values. Best results were obtained from these two combinations: temporal subtraction with amplitude thresholding, and temporal subtraction with regional-growing thresholding. To summarize, the researcher's contribution was to assess the effectiveness of various thresholding methods in the context of a three-stage approach, to help radiologists find cancerous tissue lesions in CT and MRI mammography images

Acknowledgements

Table of Contents

List of Tables

Tables

<p style="text-align:center">List of Figures</p>

Figures

Chapter 1

Introduction

Statement of the Problem and Goal

Breast cancer is the second leading cause of cancer death among women, and most women have had yearly mammography images taken to detect it (i.e., malignant growth). After a mammography image is taken, it is then reviewed by a radiologist for detection of breast cancer. After a radiologist reviews the mammography images, the images are stored over time; they are not recalled or used again by the radiologist.

The goal of this dissertation was to design an application to help radiologists and physicians better detect breast cancer. Using computer aided diagnosis (CAD) systems as tools to help in the detection of breast cancer in computed tomography (CT) mammography images, previous CT mammography images helped enhance the next series of images. If there were two CT mammogram images taken some time apart, radiologists and physicians analyzed the later CT image (B) for malignant growth. Radiologists also enhanced the data by a process using an earlier CT image (A) in a three-stage approach.

The first stage image enhancement process began with applying image registration and subtraction. This was one of the most important processes according to Miyake, Kim, Itai, and Ishikawa (2009). If misregistration occurred, this caused the subtraction artifacts to produce a false-positive result. Itai, Kim, and Ishikawa (2008) developed a new registration method with voxel-matching, a technique for removing artifacts in

temporal subtraction tomography. Previously, two types of subtraction techniques were being used for CT images: temporal subtraction and dual-energy subtraction. The researchers investigated both techniques.

The researcher accomplished the second stage by applying thresholding techniques as shown in Figure 2. Thresholding is the process of finding individual pixels in an image that are greater than a specified threshold value (Gonzalez & Woods, 2002). The researcher used eight different types of thresholding methods in this dissertation: amplitude thresholding, global thresholding, local thresholding, connected-component labeling with absolute threshold and variance threshold, threshold in histogram analysis, thresholding with K-means clustering, regional growing thresholding algorithm, and split-and-merge thresholding algorithm.

Miyake et al. (2009) used rule-based methods and artificial neural network classifiers to detect features based on shape and density in order to find lung nodules. This technique can provide a useful CAD application for feature detection of malignant growth to help radiologists and physicians detect and diagnose possible cancerous growths. Gilbert, Astley, and Gillian (2008) used CAD for screening mammography images. The researchers reported that computer-aided detection had potential cost savings and that having two readers—human and machine—had the potential to improve cancer detection rates.

This dissertation referenced eight different types of thresholding functions that were based upon the density of brightness values present in a region. Thresholding is widely used in image processing to generate binary images from gray-level images, and the researcher classified the current thresholding techniques available as either global or local

thresholding (Gonzalez & Woods, 2002). Global thresholding methods may be further broken down into point-dependent and region-dependent techniques. Global thresholding uses one threshold over the entire image to compute the binary image (Gonzalez & Woods, 2002). Typically, many features that were present in the original gray-level image were lost in the resulting binary image. A better method is to apply thresholding locally, which maximizes edge features by thresholding local regions separately (Gonzalez & Woods, 2002). Radiologists and researchers can choose these global threshold values automatically or manually, but typically chose manually when investigating the fitness of algorithms. Pixels with a gray level above the current threshold value are changed to one gray level, which is usually white (1), and the remaining pixels to another gray level, which is usually black (0), resulting in a binary image.

Algorithms such as image segmentation and image recognition operate best on binary images. Typically, the optimal threshold value is selected by maximizing a given criterion, based on the separation of an object from the background in a gray-level image or the recognition of key features in a binary image.

Segmentation-based image techniques deal primarily with those parts of an image that consist of equivalent brightness levels. These homogenous, contiguous assemblies of pixels were called regions. The researcher considered a threshold method point dependent if the threshold intensity selection was based on the gray level of each pixel.

This dissertation describes the various techniques for segmenting a group of pixels. The researcher placed emphasis on global and local thresholding techniques, which were employed for the evaluation of experiments. The researcher compared these various

thresholding methods based on their effectiveness for detecting cancerous regions. Through this dissertation, the researcher investigated methods of using CAD to assist in the accuracy of image classification and the speed in which it was performed. The resultant set of algorithms and methods should assist radiologists in the speed and accuracy of interpreting results of mammographic, gray-scale medical images.

Relevance and Significance

As the medical community looked for ways to reduce healthcare cost by investing in electronic health information systems, including electronic health records, and as more hospitals and medical offices moved their clinical environments towards computerization, centralization, and digitization, ever-increasing amounts of CT digital mammogram images were generated. Using CAD applications for feature detection of malignant growth, these historical mammogram images further enhance the interpretation of the current mammogram image and help radiologists and physicians detect and diagnose possible cancerous growths.

Why are image thresholding methods used in the first place? Clustering is the most frequently used automatic thresholding method used for classification of medical images. Clustering sorts the histogram of the image within a discrete number of classes corresponding to the number of phases perceived in an image. The gray values in the medical image are determined, and a center is determined for each class. This process repeats until a value that represents the center of the mass for each phase or class is obtained.

Thresholding is a critical step towards image understanding especially in regards to automatic analysis and image enhancement of cancerous tissue in mammography medical

image segmentation. Many current problems in image-guided surgery, therapy

evaluation, and diagnostic tools strongly benefit from the result of the analyses of

thresholding images (Maurer & Fitzpatrick, 1993). Radiologist use some popular

approaches for threshold techniques. The amplitude thresholding method manipulates

objects and backgrounds in such a way that object and background pixels have grey

levels grouped into two dominant and different modes. In local thresholding, the original

image is broken down into smaller sub-regions, and each is given a threshold

individually. Threshold-based techniques are often referred to as *histogram-based*

methods, which make decisions on local pixel information and are effective only when

the intensity levels are far outside the range of the levels in the background. K-means

clustering provides sequential splitting, initially performing dynamic splits. These

dynamic splits are then further split into smaller, more variant regions until the regions of

interest are isolated. Regional-growing methods use recursive splitting to partition the

image top-down by using the average brightness of a region. K-means algorithm is one

such region-based method used to segment the image (Tou & Gonzalez, 1974). Finding a

threshold by clustering the histogram is designed to pick the threshold in such a way that

each pixel on each side of the threshold is closer to the mean of all the pixels.

   Region-based approaches use similarity among pixels to find different regions.

Threshold methods, often used in the medical realm, partition an image into separate

regions, which ideally correspond to different real world objects. Pixels in the region are

similar to each other with respect to some characteristic property such as size, color,

intensity, or texture. The goal of thresholding is to simplify the image representation into

something that is more meaningful and easier to analyze (Miyake, Kim, Itai, & Ishikawa,

2009). This critical step enhances overall understanding of the medical-related images used in image-guided surgery, therapy evaluation, and diagnostic tools.

This critical step enhances overall understanding of the medical-related images used in image-guided surgery, therapy evaluation, and diagnostic tools. One such CAD application was image subtraction, which computed the absolute difference between two images. The first image was picked as a reference point and subtracted from the other image (Sampat, Markey, & Bovik, 2005). The subtraction of CT mammogram images that are obtained over one's lifetime could become a tool for radiologists and physicians to detect cancerous growths.

Miyake et al. (2009) examined temporal subtraction images obtained by the voxel matching and threshold techniques to find lung nodules. The authors used a filter to reduce false positives, and used a rule-based method and artificial neural network classifiers to identify lung nodules. The authors took a hard look at many thresholding techniques and used those techniques throughout the research along with temporal image subtraction. This research built on the work of Miyake et al. by investigating image registration, image subtraction methods, and thresholding methods.

Miyake et al. (2009) also discussed the use of threshold techniques based on pixel value in a temporal subtraction image. Miyake et al.'s overview, shown in Figure 1, depicts the overall scheme for this detection of lung nodules. As shown in Figure 2, this dissertation evaluated the two different techniques, temporal subtraction and dual-energy subtraction, using a rule-based method with classifiers based on decision trees and CAD design to find malignant growth in CT digital mammography.

Implementing image subtraction as an enhancement technique required the sequence

of images that are obtained over time (Gonzalez & Woods, 2002). Gonzalez and Woods

(2002) discussed image enhancement by using image subtraction, and briefly described

the main area of its application in the field of medicine where subtraction may be used for

x-ray images.

```
            ┌──────────┐
            │  Start   │
            └──────────┘
                 │
                 ▼
    ┌─────────────────────────────┐
    │   Input current and previous CT │
    │         image sets          │
    └─────────────────────────────┘
                 │
                 ▼
    ┌─────────────────────────────┐
    │     Temporal subtraction    │
    └─────────────────────────────┘
                 │
                 ▼
    ┌─────────────────────────────┐
    │  Detection of lung nodules by use of │
    │    multiple threshold technique │
    └─────────────────────────────┘
                 │
                 ▼
    ┌─────────────────────────────┐
    │  Elimination of false positives by use │
    │   of selective enhancement filter, │
    │    rule-based method with ANN │
    │          classifier         │
    └─────────────────────────────┘
                 │
                 ▼
            ┌──────────┐
            │   End    │
            └──────────┘
```
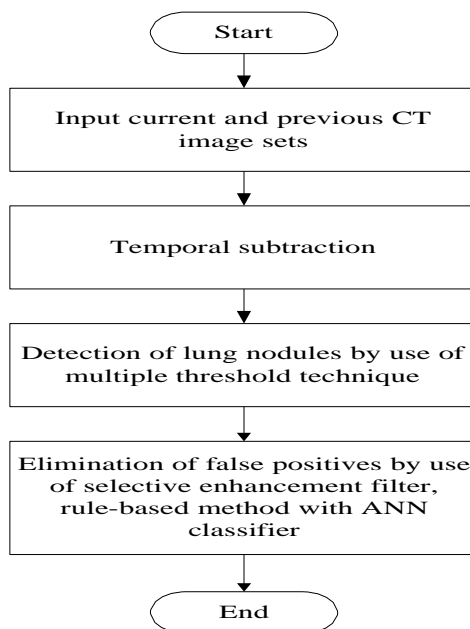
Figure 1. Overall scheme for detection of lung nodules (Miyake, Kim, Itai, & Ishikawa,

2009).

One other technique for medical image enhancement is image segmentation by

thresholding (Gonzalez & Woods, 2002). CT mammogram images were segmented by

variance blocks with lower variances that were segmented from blocks with higher

variances. A CT mammogram image was composed of objects and backgrounds in such

a way that object and background pixels had gray levels grouped into dominant and

different modes (Singh & Al-Mansoori, 2000).

Several researchers investigated a number of region-growing techniques (Kunt et al.,

1985). Kunt, Ikonomopoulos, and Kocher (1985) started to use region-growing

techniques with feeble uniformity predicates. After preprocessing the image, Kunt et al. assigned pixels to regions according to a global threshold distribution, such as 0–9, 10–19, 20–29, etc. (out of a full, dynamic range of 0–255 brightness levels). This led to an enormous number of false regions in which regions were split where no contour existed. In one of the testing images, the low-frequency brightness change that occurred in the largely uniform sky provoked it to split into eight distinct large regions, and many small regions on the borders (Kunt, Ikonomopoulos, & Kocher, 1985).

Kunt et al. (1985) stated that using a more complex uniformity predicate would simply take too much time (i.e., CPU resources). Instead, the researchers adopted an adaptive, quadtree-split technique. The split process followed that of the quadtree, generating geometrically equal sub-regions. The choice of splitting regions was controlled by information obtained by a boundary detection algorithm (a region is split if a strong contour is detected within the image). These contour-texture oriented techniques attempted to describe an image in terms of contour and texture.

Miyake et al.'s (2009) process had three steps: registration and subtraction, thresholding, and elimination of false positives using artificial neural network (ANN) classifiers to identify lung nodules. The researchers had mixed success. Interestingly, Miyake et al. used temporal subtraction, whereas Carton, Lindman, Ullberg, Francke, and Maidment (2007) used a dual-energy subtraction technique to find breast cancer, as they found registration between temporal images to be unreliable. This dissertation researcher used the Miyake et al. process for stage one. For the dissertation, the researcher added voxel-matching for the registration and focused on the subtraction and thresholding steps, keeping the first stage's input current, and previous CT image sets fixed, and the

researcher utilized a decision tree in the last stage.

Image segmentation is the process of partitioning an image into regions of uniform brightness (Pavlidis, 1980). Pavlidis (1980) investigated variance thresholding, linear approximation, clustering by K-means, connected-component labeling, and regional aggregation techniques as methods for further enhancement of the image after the image had been subtracted, providing the absolute value of the image. Miyake et al. (2009) talked about using a rule-based method, but first the researchers used three image segmentation methods, such as thresholding, for the detection of lung nodules that were candidates for cancer. Region-splitting algorithms were predominant in the literature of segmentation-based images in both Kunt's and Pavlidis' work.

The researcher used the following evaluations for splitting behavior in this study. The relationship between segmentation and region splitting was that region splitting was a segmentation approach. Region splitting recursively partitioned the image top-down by using the average brightness of a region, the average of the minimum and maximum brightness present in the region, and K-means clustering where the number of regions "K" was fixed in advance. The researcher may have achieved shape simplification by eliminating the very large regions in an image. Therefore, the researcher initially expected a sequential splitting to perform dynamic splits, and to follow with splitting of smaller, more variant regions. The researcher will elaborate on this process in the chapter on methodology. In this dissertation some of the experiments evaluated thresholding with first-order statistics, thresholding with histogram valleys, and splitting by K-means clustering. To quantify performance of the algorithms, experiments using each algorithm were performed.
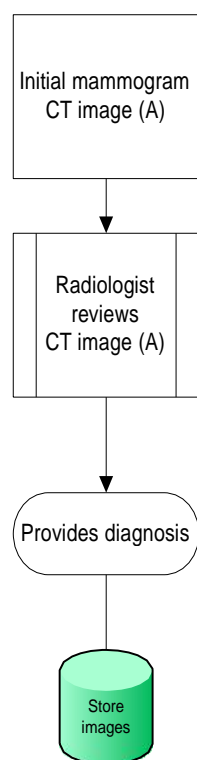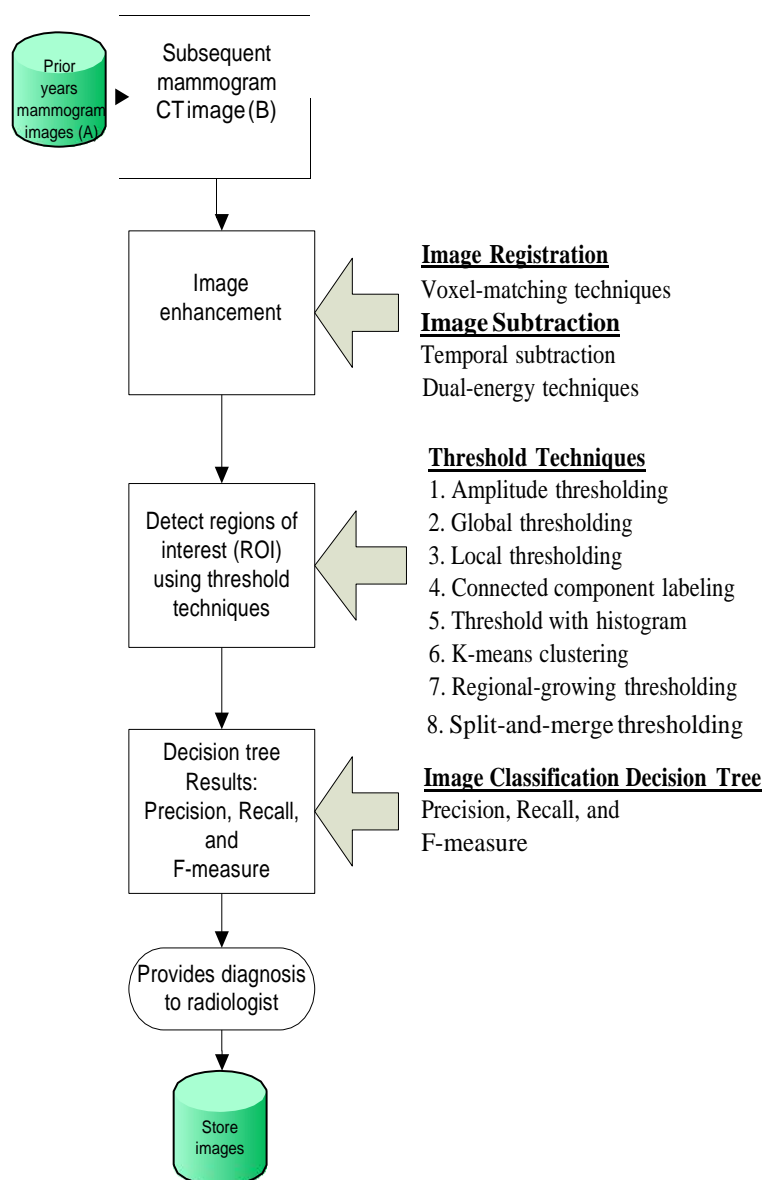
Figure 2. Proposed process.

Barriers and Issues

In this study, splitting schemes, although more time-consuming, had the advantage of

being able to make more reliable estimates of predicates over a large group of pixels

(Pavlidis, 1980). CT and MRI images were three-dimensional, taken from different angles. Since contrast-enhanced CT and MRI images carried fine and vital details, the amount of signal-to-noise ratio and distortion introduced to the image needed to be examined as a potential issue.

The current framework for this dissertation was adapted from the work of Miyake et al. (2009). One of the key issues involved obtaining the hardware and computer software for this framework. In addition, the researcher needed to create a software framework to evaluate different the thresholding algorithms used in the evaluation section.

Obtaining appropriate sets of images for comparison in the research method section of this dissertation— i.e., temporal subtraction versus dual-energy subtraction—was difficult, as dual-energy subtraction for contrast-enhanced, digital breast tomosynthesis is fairly new technology.

Elements, Hypotheses, Theories, and Research Questions to Be Investigated

In this study, the researcher categorized image segmentation into three classes (Rosenfeld & Kak, 1982). The first was segmentation by global characteristics. Its fundamental idea encompassed the possible existence of differences of intensity between objects and their backgrounds, within global and local thresholding, in which the researcher then determined the intensity. Consequently, all the pixels with a value less than the threshold belonged to one class and the remaining points belonged to another class. The second method was detecting edges and lines. The principle of this method was to consider edge and line boundaries between objects. By detecting boundaries, objects were then identified. Thirdly, extracting regions was an approach that searched for regions according to some criteria such as uniformity of intensity and conformity to

certain geometric shapes. The following questions were addressed in the experimentation

section of this dissertation:

Voxel-matching for registration:

      1. Was it effective for CT and MRI mammogram images?

Subtraction; Type 1 (temporal) Type 2 (dual-energy):

1. Did one type produce better initial results for CT and MRI mammogram
   images?

2. What impacts did Type 1 and Type 2 have on the results?

3. What were the pros and cons of each type?

Thresholding:

1. Was it possible to extract the objects from the background using thresholding?

2. What was best for finding malignant growth in CT and MRI mammogram
   images, global thresholding methods or local thresholding methods?

3. How did considered thresholding methods compare in finding malignant
   growth in CT and MRI mammogram images?

4. Did the thresholding methods identify with ROI in the original classification?

5. Did the size of the malignant growth matter in any of the thresholding
   methods used?

Based on preliminary research, the researcher investigated the following hypotheses for

validity:

1. Voxel-matching for registration. CT and MRI images worked with region
   mapping, for example, Linh and Linh (2010) talked about the voxel-matching
   process with CT and MRI images, and reported that in most cases it

successfully discriminated regions. This dissertation expected that with CT and MRI images, the ROI could be easily distinguished from the tissue region by using threshold algorithms.

2. Image subtraction. Miyake et al. (2009) successfully used temporal subtraction for the detection of lung nodules. Carton et al. (2007) completed studies to find cancer using dual-energy. In addition to investigating both of those techniques, previous researchers indicated that dual-energy may have provided the best solution (Carton et al., 2007).

3. Thresholding. The researcher investigated both global and local thresholding methods. Local thresholding initially looked to be a better solution, because it maximized edge features by thresholding local regions separately. Regional growing looked to be the most effective way for thresholding CT and MRI medical images, because region-growing techniques did not pose any restrictions on maximum region size or shape. A combination of local thresholding with regional-growing techniques could have provided the best overall classification.

Limitations and Delimitations

Digital images obtained by CT and MRI scans were discrete both spatially and in brightness (Gonzalez & Woods, 2002). It was denoted by the function f (i, j), where the arguments i and j are spatial coordinates that specified the location. The value of the function f is the intensity at location i, j, which is called the grayscale. The researcher confined this dissertation to grayscale CT and MRI mammography medical images, which were a sequence of frames acquired by a gamma camera. Specifically, a single

pixel was represented by 8 bits; i.e., the dynamic range of brightness from 0 (black) to 255 (white).

Definition of Terms

An Artificial Neural Network (ANN): an artificial system that adaptively changes its structure based on external or internal information that flows through the network.

Back-Propagation Network (BPN): a field of pattern classification created by mapping input patterns into one or several categories.

Benign tumors: non-cancerous tumors that grow larger but do not spread to other parts of the body.

Classifiers in decision trees classification: often measured by precision, recall, and f-measure. An artificial neural network refers to a type of learner, loosely inspired by the interconnected nature of biological neurons.

Computed Tomography (CT) *mammogram images*: medical images used in diagnostic radiology, including digital radiographs and nuclear magnetic resonance, created from a machine designed specifically to x-ray breast tissue. In CT images, intensity is proportional to the absorption rate of an x-ray beam in the object being irradiated. Contrast agent can be used to highlight anatomical structures or tissues.

Computer Aided Diagnosis (CAD): a tool that assists radiologists and physicians to interpret diagnostic information.

Contrast Agent (CA): an iodine substance used to enhance the visibility of computed tomography x-ray imaging.

Decision Trees: graphs in a shape resembling a tree where each branch node represents a choice between a number of alternatives, and each leaf node represents a

classification.

Histogram: constructed for the image intensity values called a gray-level range.

K-means clustering: represents a group of pixels with homogeneous gray levels.

Magnetic Resonance Imaging (MRI): intensity is proportional to the Nuclear Magnetic Resonance (NMR) signal intensity of the contents of the corresponding volume element or voxel of the object being imaged.

Malignancy: several main types of malignancy exist. Carcinoma is a malignancy that begins in the tissues that line or cover internal organs.

Regions of Interest (ROI): all points contained within boundaries of an area of interest that can be defined on an image.

A temporal subtraction image: obtained by subtraction of a previous image from a current one; can be used for enhancing interval changes (such as formation of new lesions and changes in existing abnormalities) on medical images by the removal of most normal structures.

Tomosynthesis mammogram images: three-dimensional digital images from different angles. Each slice is composed of two-dimensional elements called pixels. A tomographic image is composed of several picture elements called pixels. The intensity of a pixel depends on the modality used in the image.

A voxel: a data-set that is composed of three-dimensional elements.

Summary

In 1973, G. N. Hounsfield (a senior research scientist in Middlesex, England) invented a revolutionary imaging technique that he called computed axial transverse scanning. Hounsfield presented a cross-sectional image of the head that revealed the tissues inside

the brain as separate structures of gray matter, white matter, CSF (Cerebrospinal Fluid), and bone. This new imaging method allowed pathologic processes such as blood clots, tumors, and infarcts to be easily seen. Anatomical structures inside the human body that had never been imaged before could now be visualized. During the 41-year period since its discovery, this imaging modality has completely revolutionized the practice of medicine (Gonzalez & Woods, 2002). At the beginning, the CT scanners only could image the head, but now the scanners also have primary roles in diagnosing disorders of the chest, abdomen, and pelvis (Gonzalez & Woods, 2002). The original scanners took several minutes to acquire a single slice through the brain. The newest scanners now can image the entire body in one to two minutes.

The cue for implementing CT and MRI mammogram image subtraction as an enhancement technique for the sequence of medical images was obtained from Gonzalez and Woods (2002). The researchers discussed image enhancement by image subtraction and image segmentation (Gonzalez & Woods, 2002). *Image segmentation* is the process of dividing an image into units that are homogeneous with respect to one or more characteristics (Pavlidis, 1982). This included the processes of identifying partitions, which possess distinct features, such as gray levels or texture (Pal & Pal, 1993). The level to which subdivision was made depends on the problem to be solved. That is, segmentation stops when the object of interest in an application has been isolated (Jain, 2002). Hence, image segmentation was regarded as the first and most critical step for much of the image analysis (Pal & Pal, 1993). Splitting an image implied starting at the top of the structure and traversing downward. On the other hand, the merging approach started at the pixel level and merged upward in the hierarchy of regions.

This dissertation considered pairs of two CT and MRI mammogram images taken some time apart to find the differences. Using voxel-matching, temporal and dual-energy subtraction techniques provided the absolute value of the two images. This enhanced the image by bringing out the finer details. Karellas, Vedantham, and Lewin (2009) offered a survey paper from the University of Massachusetts Medical School, and talked about temporal subtraction and dual-energy subtraction techniques as two methods for increasing visibility in pilot studies. Miyake et al. (2009) used CAD design to detect lung nodules. The researcher of this study implemented the same type of CAD design to detect malignant growth in CT and MRI mammography images by using different thresholding techniques. Miyake et al. (2009) mainly looked at histograms as a thresholding technique and only looked at temporal subtraction with voxel-matching. For this dissertation, the researcher used voxel-matching for the registration and used both temporal subtraction and dual-energy subtraction.

Miyake et al. (2009) used thresholding techniques based on a pixel value. This dissertation also used thresholding techniques such as amplitude thresholding, first-order statistics of the accuracy of segmentation via thresholding, global thresholding, connected-component labeling with absolute threshold, connected-component labeling with variance threshold, threshold in histogram analysis, and thresholding with K-means clustering algorithm. An image was split into regions of nearly constant brightness, and then those regions were split again, and so on until the desired outcome was achieved. Previous researchers investigated variance thresholds, clustering by K-means, amplitude variance thresholding by means of connected-component labeling, and region aggregation techniques as methods for defining suitable regions and sub-regions for

splitting (Gonzalez & Woods, 2002).

The researcher of this dissertation enhanced the image by using image temporal subtraction, and dual-energy method subtraction techniques. The global and local algorithms that were introduced posed no restrictions on region shape or size, and were tailored to allow constraints to be applied and approximated with ease (via tree-structuring).

The work of Miyake et al. (2009) was similar, but differed technically in that the researchers only used temporal subtraction and neural networks, whereas this dissertation looks at two different types of image subtraction: temporal and dual-energy. Through this dissertation, the researcher looked at eight different thresholding methods and uses decision trees. In addition, Miyake et al. (2009) looked for lung nodules, whereas this dissertation was looking for malignant growth in CT and MRI mammogram images. The researcher of this dissertation looked at both temporal subtraction and dual-energy subtraction techniques with CT and MRI mammogram images taken some time apart. Regional thresholding algorithms, such as the one proposed, were generally noise-tolerant and thus less prone to information loss from image pre-processing (noise filtering). The splitting threshold methods could be derived from:

1. Amplitude thresholding,

2. Global thresholding methods,

3. Local thresholding methods,

4. Connected-component labeling with absolute and variance thresholding,

5. Threshold with histogram analysis,

6. K-means clustering where "K" is a variable,

7. Regional growing thresholding algorithms, and

8. Split-and-merge thresholding algorithms.

The researcher used the following to analyze the results:

1. Accuracy of segmentation. How well did the thresholding method identify the correct region of interest as compared to the original classification?

2. Precision of classification. How well did the resultant classification agree with pre-classified images?

3. Analysis of results. The test set is derived from applying a thresholding method to a single image against the training set. Each image type has its own decision tree. Which thresholding and subtraction method combinations yielded the best precision, recall, and f-measure values?

Chapter 2

Review of the Literature

The Theory and Research Literature Specific to the Topic

The basic principle behind CT images is that the internal structure of an object can be reconstructed from multiple projections of the object (Gonzalez & Woods, 2002). A patient lies on the table within the CT gantry, which is shaped like a giant donut. During each slice acquisition, an x-ray tube circling the patient produces an x-ray beam that passes through the patient, and is absorbed by a ring of detectors surrounding the patient (Gonzalez & Woods, 2002). The intensity of the x-ray beam that reaches the detectors is dependent on the absorption characteristics of the tissues it passes through (Gonzalez & Woods, 2002). Since the beam is moving around the patient, each tissue is exposed from multiple directions. Using Fourier analysis, the computer uses the information obtained from the different amounts of x-ray absorption to reconstruct the density and position of the different structures contained within each image slice (Hornak, 1999).

One of the most frequent problems found in computer vision and image analysis is that of performing an element-by-element or region-by-region comparison of one observed image with a near identical reference image (Miyake et al., 2009). This image registration takes two images of the same object that may be acquired under different conditions (Miyake et al., 2009). For example, images taken from different times; hence, the images to be compared may have differed in intensity and in geometry. To compute a correct comparison, one image must be aligned with the other in such a way that all

corresponding points or regions match. Miyake et al. (2009) mentioned registration between images as one of the most important parts of analysis. Misregistration could cause subtraction of artifacts and false positives in CT images.

Implementing image registration and subtraction as an enhancement technique requires an image sequence that is obtained over time (Gonzalez & Woods, 2002). Miyake et al. (2009) discussed image enhancement by using temporal image subtraction. The researchers briefly described the main area of its application, in the field of medicine, where subtraction may be used for CT images. Carton et al. (2007) used a dual-energy subtraction technique for contrast-enhanced breast tomosynthesis to find breast cancer. The subtraction of CT mammogram images that are obtained over one's lifetime could lead to a CAD tool for radiologists and physicians to find breast cancer.

The size of regions with equivalent brightness in digital areas of coherence tend to be large in CT images, which is better because the CAD tool can encode regions to be dramatically reduced in terms of data volume. However, before these regions could be encoded in this study, the researcher had to identify the regions, and approach the problem of locating these regions. One option was image segmentation, the process of subdividing an image into regions of uniform brightness (Pavlidis, 1980). By applying proper segmentation techniques in this study, the researcher extracted featureless regions or the background from image details. Miyake et al. (2009) found that inhomogeneous background artifacts could cause problems and could be therefore eliminated. The authors used a background trend correction that was applied on the original and subsequent images, which was a morphological filter before they used temporal subtraction (Miyake et al., 2009). Wu (1992) recognized the advantages of free

segmentation, fewer regions, and better conformance to the real corners, but discarded the idea because "natural segmentation can have shapes too complex to be encoded compactly" and "it is computationally too expensive," the latter referring to Kunt's work (Kunt et al., 1985).

However, Kunt et al. (1985) used pre- and post-processing techniques on contours created by their region-growing algorithm. Kunt et al. (1985) examined the property used for describing whether a pixel should be appended to a region, which is membership in a grey-level interval (e.g., 0–9, 10–19, etc.). This property was too simple according to Murat Kunt (1985), who stated that it would require hours of CPU time to execute. Instead, Kunt, like Xiaolin Wu, resorted to a quadtree-based, split-and-merge technique (Kunt et al., 1985). Several authors had investigated image segmentation with different solutions, including Kunt et al. (1985), and started to use region-growing techniques with feeble uniformity predicates. After preprocessing the image, Kunt et al. assigned pixels to regions according to a global threshold distribution, e.g., 0–9, 10–19, 20–29, etc. (out of a full dynamic range of 0–255 brightness levels). This led to an enormous number of false regions; regions were split where no contour existed. In one of Kunt et al.'s testing images, the low-frequency brightness change that occurred in the largely uniform sky provoked it to be split into eight distinct, large regions and many small regions on the borders.

In Runtime analysis, CPU timing is affected by various factors, many of which are related to the threshold algorithm itself, but it supplies the amount of average time over multiple images for each of the four threshold methods used in this study. The second question in this study, regarding the test data, is true that *constant time algorithms*, by

definition, take the same amount of time on different inputs. The researcher will address these values in the appendix.

Several authors investigated image segmentation with different solutions (Saidin et al., 2010). Contour-texture oriented techniques attempt to describe an image in terms of contour and texture (Saidin et al., 2010). According to Saidin et al. (2010), seed-based, regional-growing techniques started to take off in the area of mammogram images in both magnetic resonance imaging MRI and CT mammogram images.

The final step in Miyake et al.'s (2009) work was implementing artificial neural network (ANN) classifiers to reduce the number of false positives. The researchers used statistical features based on six shape features and four density features. Besides Miyake et al., many authors employed ANN classifiers for the detection of breast cancer, some achieving 93.6% accuracy (Salma, Abdelhalim, & Zeid, 2013). In addition, neural networks comprise a class of computing paradigms that recently spawned a great deal of research (Salma et al., 2013). According to Miyake et al. (2009), one widely used neural network paradigm is the multi-layer perception, employing back-propagation of errors learning, often called a Back-Propagation Network (BPN). BPNs had proven to be very useful in the field of pattern classification by mapping input patterns into one of several categories. Rather than being specifically programmed, BPNs "learned" this mapping by exposure to a training set—a collection of input pattern samples matched with their corresponding output classification (Miyake et al., 2009). The proper construction of this training set was crucial to successful training of a BPN. Miyake et al. (2009) used an ANN with a three-layer-based, back-propagation algorithm to remove false positives. Nodule rule-based method and artificial neural network classifiers can detect features

based on shape and density.

One criteria to be met for proper construction of a training set is that each of the classes is adequately represented. A class that is insufficiently represented in the training data may not be learned as completely or correctly, impairing the network's discrimination ability. This is because of the implicit setting of prior probabilities, which results from unequal sample sizes. The degree of impairment is a function of, among other factors, the relative number of samples of each class uses for training (Salma et al., 2013).

Image Enhancement

The principal objective of enhancement techniques was to process mammography medical images so that the results were more suitable than the original image (Gonzalez & Woods, 2002). The enhancement process may not have increased the inherent content in the image, but it did increase the dynamic range of the chosen features so that it was more useful for display (Jain, 2002). Different techniques of enhancing were followed in different applications; image subtraction may be useful for enhancing x-ray images.

The past approaches for image enhancement fell into two broad categories: spatial domain methods and frequency domain methods (Gonzalez & Woods, 2002). The spatial domain method referred to the mammography image coordinate itself, and researchers based approaches in this category on direct manipulations of pixels in an image (Gonzalez & Woods, 2002). Frequency domain processing techniques were based on modifying the spectral transforms of a mammography image.

Voxel Image Registration

In the literature of image registration, researchers proposed a wide variety of methods

for medical or non-medical applications. In general, any registration method produced a set of equations that transformed the coordinates of each point in one image into the coordinates of the corresponding point in the other image. Registration methods may be classified according to the various criteria. Researchers limited most of the classification approaches to the retrospective methods, i.e., methods that register data after acquisition.

Maurer and Fitzpatrick (1993) defined *registration* as the "determination of a one-to-one mapping between the coordinates in one space and those in another such that points in the two spaces that correspond to the same anatomical point are mapped to each other" (Maurer & Fitzpatrick, 1993) (see Figure 3). Rigid body motion could be decomposed into a rotation and a translation. Researchers can easily compute scaling from the imaging specifications associated with each medical modality (pixel size, distance between slices). Point-based registration methods, also referred to as *point-fitting methods*, require the determination of the coordinates of some corresponding points (feature points) in different images, and the estimation of the geometric transformation using these corresponding points (Gonzalez & Woods, 2002). The feature points were either intrinsic, extrinsic, or a combination of both. Intrinsic points, also called *fiducial points*, derive from specific image properties, usually anatomical landmarks (Gonzalez & Woods, 2002). Extrinsic points derive from artificially applied markers called *fiducial markers* (Gonzalez & Woods, 2002).

$$I_1(x,y) \longrightarrow \boxed{\text{Mapping}} \longrightarrow I_2(x,y)$$
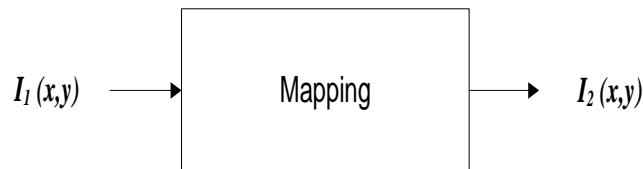
Figure 3. Mapped images.

Zitová and Flusser (2003) stated that *image registration* was a process that

overlaid two or more images of the same scene taken at different times. Linh and

Linh (2010) defined image registration as the process of using different sets of data of

the same object, which came from different modalities (CT, MRI, SPECT, PET, etc.).

The image registration process geometrically aligned the two images. Image

registration was a process known for transforming different data sets onto one

coordinate system. Researchers believe image registration is the fundamental step in

combining computer tomography (CT) image analysis (Zitová & Flusser, 2003).

Figure 4 shows the registration methods consisting of the following four components.

The registration process, according to Zitová and Flusser (2003), includes:

1) Feature detection extracted informational measures from the images that were

   used for matching. For instance, consider the problem of registering two

   images of the same building taken at different times, shown in Figure 4. A

   registration approach of the images might be as follows: Feature detection was

   first performed; relevant and distinguishing objects (closed-boundary regions,

   edges, contours, line intersections, corners, etc.) were automatically detected;

   then for further processing these features were represented by their point

   representatives (centers of gravity, line endings, distinctive points), which

   Zitova and Flusser (2003) called control points (CPs).

2) Feature matching decided how to choose the next feature from the space to be

   tested in the search for the optimal change. With feature matching, the

   association between the features detected within the sensed image and those

   associated within the reference image were established. A mixture of

   descriptors and likeness, in company with spatial relationships, along with the

features, were used for this purpose.

3) Transform model estimations were the class of transformations capable of aligning the sensed images. These types of parameters, called mapping functions, aligned the sensed image with the reference image. These parameters and the mapping functions could be computed by means of the recognized feature correspondence.

4) Image resampling and transformation determined the relation of each test search continuously, in accordance with the search strategy, until a transformation was found on the image that comparison measurement was satisfactory. Then the sensed image was transformed by means of mapping functions. The image values in non-integer coordinates were then computed by the suitable method.
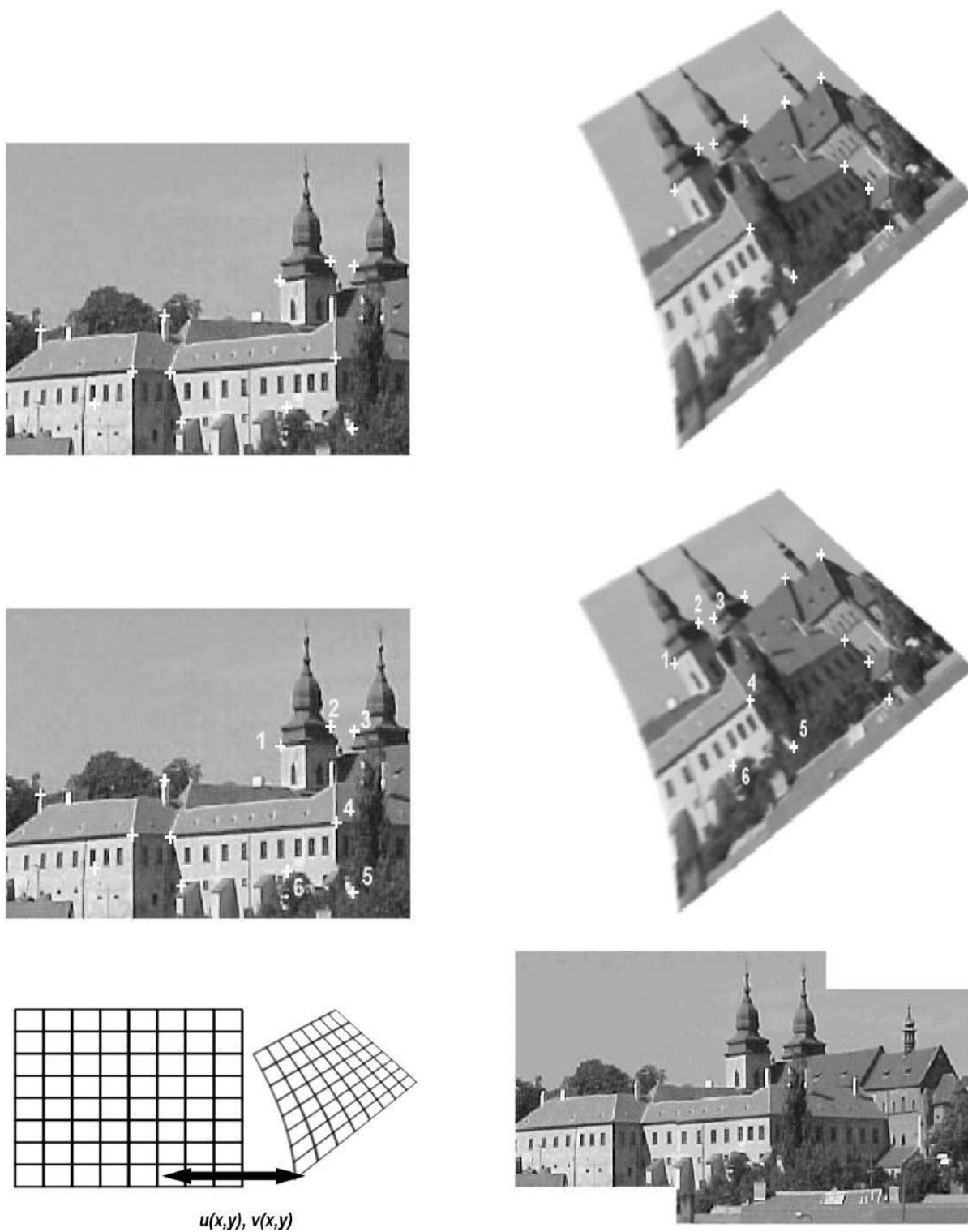
Figure 4. Registration process (Zitová & Flusser, 2003).

Magnetic resonance imaging (MRI), computed tomography (CT) images, and all other tomographic modalities are imaging modalities for producing images of a slice through the human body (Tomography, from the Greek *tomos*, means slice). This form of imaging is in some respects equivalent to cutting off the anatomy above the slice and below the slice. Each slice is a thickness (Thk), and the slice is said to be composed with a matrix of volume elements or voxels (see Figure 5). The tomographic image is composed of several picture elements called pixels. The intensity of a pixel depends on the modality used; for example, MRI is a medical imaging technique used to produce high-quality images of the inside of the human body. The technique is called *magnetic resonance imaging*, rather than nuclear magnetic resonance imaging (NMRI), because of the negative connotations associated with the word nuclear in the late 1970s (Hornak, 1999). Magnetic resonance imaging started out as a tomographic imaging technique, i.e., to produce an image of the nuclear magnetic resonance (NMR) signal in a thin slice through the human body (Hornak, 1999). Since then, magnetic resonance imaging advanced beyond a tomographic imaging technique to a volume imaging technique.
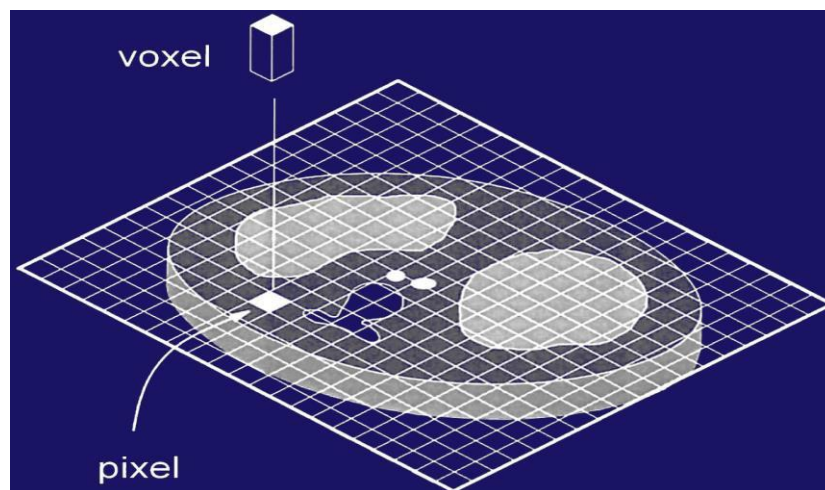
Figure 5. Dataset composed of three-dimensional elements (Hornak, 1999).

In the literature, researchers have proposed a wide variety of image registration methods for medical and non-medical applications (Hornak, 1999). In general, any registration method produces a set of equations that transforms the coordinates of each point in one image into the coordinates of the corresponding point in the other image. Registration of medical images has been the subject of extensive study in the medical imaging field. Medical image registration is important for several reasons; one such reason is intra-subject, intra-modulator registration used to monitor treatment. One example of this is the need to allow accurate comparison of lesion size and position during treatment over time. This process allows precise comparison, in time, of the evolution of a lesion.

Image Subtraction

The Burt and Adelson (1983) method reduced the variations in illumination over the entire image by subtracting the low-frequency, grey-level image from the original grey-level image. This method was known as the image subtraction method (Hornak, 1999). The technique segmented the original grey-level image into regions using the low-frequency, grey-level image to determine large regional grey-level areas. The researcher then gave each region a threshold individually by using a global thresholding technique (Hornak, 1999).

In medical image applications it was desirable to compare two images. A simple but powerful method was to compare two images and subtract them. This image subtraction was a type of point-processing technique in which the pixels in the image were manipulated to get the desired enhancement. Itai, Kim, Ishikawa, Katsuragawa, and Doi (2008) performed image subtraction by first removing artifacts, and then by subtraction

of a previous image from a current one. In image subtraction, the absolute difference

between two images f(x,y) and h(x,y) (Jain, 2002) was expressed as

$g(x, y) = |f(x, y) - h(x, y)|$, obtained by computing the absolute difference between all

pairs of corresponding pixels in f and h. Processing medical images was an important

area of digital image enhancement (Gonzalez & Woods, 2002); it was a powerful

technique that increased the dynamic range of required features and was very helpful for

physicians in diagnoses. The image subtraction method could remove the local change in

illumination over a grey-level image. By using a low-pass, spatial filtering on the grey-

level image, Hornak (1999) created a new image with the overall contrast and average

intensity of the original image. After subtracting the new image from the original image,

the researcher removed the regional variations in illumination (Hornak, 1999). The

researcher could then threshold the resulting image using a global thresholding technique.

   Miyake et al. (2009) and Carton et al. (2007) used image subtraction. Carton et al.

(2007) stated that dual-energy, contrast-enhanced (DECE) subtraction technique

tomography images were produced by enhancement. Miyake et al. (2009) used temporal

subtraction images obtained by the voxel-matching technique to enhance images with

lung nodules. In addition, Carton et al. (2007) emphasized iodine contrast in the breast.

In the present analysis, the researcher applied subtraction for high-energy and low-energy

tomography and looked at dual-energy subtraction in addition to temporal subtraction.

The following equation from Carton et al. (2007) uses weighted subtraction that is

applied to the logarithm of the high- and low-energy tomographic images:

$$I_{DE}(x, y) = ln(I_H(x, y)) - w_i ln(I_L(x, y))$$

where I_{DE} denoted the "iodine-enhanced" DECE tomography image and the subscripts L

and H designated the low- and high-energy tomography images. The tissue cancellation

factor, w_i, acted as a weighting factor that maximized the signal difference to noise ratio

(SDNR) between the iodine contrast agent and the structured breast, and that minimized

the variance in the structured breast tissue image Carton et al. (2007).

Thresholding

A thresholding method is considered point dependent if the threshold selection

depends only on the grey-level of each pixel; for some features (e.g., contours) point

dependency does not depend. Otsu (1979) used these methods in applications such as

cell analysis. In the modal (optimum thresholding) method, Prewitt and Mendelsohn

(1966) used histograms of grey-level to make threshold selection automatic. Figure 6

depicts multiple peaks, representing both the object and background of the image. By

choosing the threshold at the bottom of the valley, the object was separated from the

background. When multiple valleys are in a histogram of an image, the peaks indicate

several large sets of pixels with the same gray-level value.
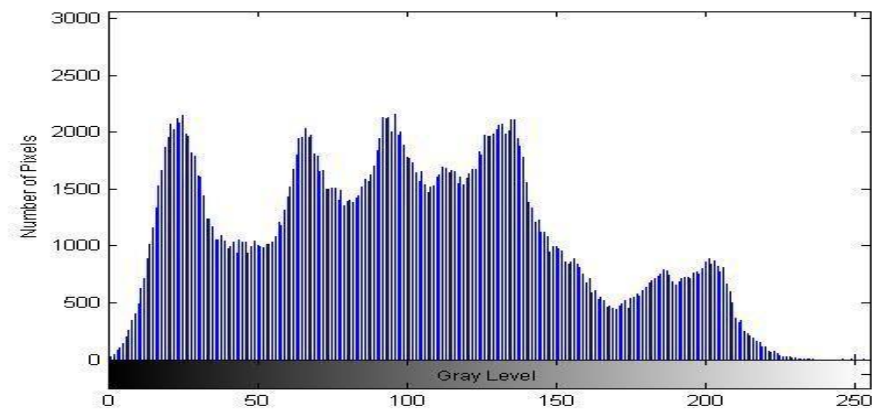
Figure 6. Matlab histogram.

One method of global thresholding proposed by Otsu (1979) was based on

discriminate analysis. It separated the image into two classes above and below the

threshold. Otsu selected a threshold automatically based on the discriminate criterion, and the goal was to select a threshold value that maximized the similarity of grey level in the two classes. Typewritten characters and texture images demonstrated the results of this method. Otsu's method also was extended to multi-thresholding methods in the case of an image histogram. To demonstrate multi-thresholding, Otsu used an image of cells with a histogram. Through this method, Otsu was able to select two threshold values that successfully separated the cell from the background and the nucleus from the cytoplasm.

Brink (1991) proposed a method similar to Otsu (1979) in which the correlation between the original and threshold images was evaluated and maximized by iteration. Brink's method was demonstrated on two different images: an image of a sample of handwriting representing an image with a unimodal histogram, and an image of bacteria colonies that was used to represent an image with a bimodal histogram. Brink's method was able to select valid threshold values for the two cases.

Other global thresholding methods included entropic, moment-preserving, and minimum-error methods (Pun, 1981; Tsai, 1985). Entropic thresholding methods use information theory to obtain the optimal threshold. Pun (1981) proposed an entropic method that used an anisotropy coefficient α in determining the threshold value. Pun used the value of the anisotropy coefficient to classify the shape of the image's histogram. The shape of the histogram determined the final selected threshold value.

Tsai (1985) proposed a moment-preserving method. The threshold was selected such that the grey-level moments of the input image were preserved in the output image. Tsai demonstrated this on images with bimodal, trimodal, and quadmodal histograms. The researcher gave the images threshold values according to the number of modes of the

histogram: one, two, or three thresholds were used for bimodal, trimodal, and quadmodal

histograms, respectively (Tsai, 1985). Seven different images were used to demonstrate

the algorithm, and Tsai's method yielded reasonable results for the images Tsai

presented.

In first-order statistics, the accuracy of segmentation via thresholding depended almost

entirely on the selection of the value for the threshold (Chen & Chen, 2006). Good

threshold values were estimated by using statistical information about the image. The

average, maximum, minimum, median, and mode of the pixels brightness were examples

of these (Verbeek, Vrooman, & Vliet, 1988).

Pavlidis (1980) defined a uniformity predicate: Let X denote the grid of sample pixels

of a picture i.e., the set of pairs

$$(i,j)\, i=1,2...,N\, , j=1,2...,M\, .$$

If Y is a nonempty subset of X, uniformity predicate P(Y) assigns the value of true or

false to Y depending on the properties of the brightness matrix f(i,j) for the points of

$(i,j) \in Y$. Also, a uniformity predicate should satisfy this constraint: if Z is a nonempty

subset of Y, then P(Y) = TRUE implies that P(Z) = TRUE. A uniformity predicate can

be defined as in the following Table 1 contains examples of uniformity predicates.

Table 1

Example Uniformity Predicates (Pavlidis, 1980)

P_1 (Y)	If the brightness value at any two points of Y is the same.
P_2 (Y)	If the brightness values at any two points of Y do not differ by more than a given amount e.
P_3 (Y)	If the brightness value at any point of Y does not differ by more than a given amount from the average value of f(i, j) taken over a set P_3 containing Y.
P_4 (Y)	If the maximum of the brightness value over Y does not exceed a given value.

A segmentation of the grid X for a uniformity predicate P was a partition of X into disjoint nonempty subsets $R_1, R_2 ... R_n$ such that (Pavlidis, 1980):

$$\bigcup_{i=1}^{n} R_i = X \ \ where \ R_i \ \cap \ R_j = \phi, \ \ \forall i \neq j$$

$$R_i \ is \ connected \ , \ \ \forall i$$

$$P(R_i) = TRUE, \ \ \forall i$$

$$P(R_i \cup R_j) = FALSE, \ \ \ \forall R_i \, adjacent \ to \ R_j$$

From this point on, the term *region* will indicate a set of pixels with the above properties. Stated differently, any subset of a connected region R was adjacent to the set containing the remaining points in R (Pavlidis, 1980).

Amplitude Thresholding

One technique for image segmentation was amplitude thresholding (Gonzalez & Woods, 2002). According to Gonzalez and Woods (2002), in some cases an image, f(x,y), is composed of objects and backgrounds in such a way that object and background pixels have grey levels grouped into two dominant and different modes. In these cases, it is possible to extract the objects from the background using a threshold that separates these modes. Image thresholding operations followed the general relation for a threshold T (Gonzalez & Woods, 2002):

$$g(x, y) = T(x, y, p(x, y), f(x, y))$$

where x, y are the current point's coordinates, p(x,y) is a local function (e.g., the average of the eight neighboring pixels), and f(x,y) is the grey level of the current point. If T depends solely upon f(x,y), then it is called a global threshold. If T depends on p(x,y) in

addition to f(x,y), it is a local threshold. Gonzalez and Woods found T to be dynamic if it depends on all four of the components.

Global and Local Thresholding

The simplest version of global thresholding techniques is bi-level thresholding. A researcher can compare every pixel brightness value with a single preset value, and then assigned it to one of two categories (Pavlidis, 1980). Given n pixels, the thresholding algorithm is O(n), rendering a fast image segmentation solution. According to Pavlidis (1980), thresholding is an image matrix f(x,y) with a single threshold T generated a binary matrix v(i,j), defined by the relation:

$$v(i,j) = \begin{cases} 0, & if \ f \ (i,j) < T \\ 1, & if \ f \ (i,j) \geq T \end{cases}$$

This method was ideal for application on simple foreground-background images, but this was hardly the case in natural scenes. When an image had different surface characteristics, or simply more than two regions with different brightness functions, multilevel thresholding provided a better classification (Pal & Pal, 1993). Thresholding with m was generalized by (Pavlidis, 1980):

$$v(i,j) = k \ if \ T_{k-1} \leq f(i,j) < T_k, \ k = 1,...,m$$

where $T_o, T_1,..., T_m$ was an ordered list of threshold values, where $T_0=0$, $T_m=256$, and $T_{k-1} < T_k$ for k=1,…,m.

Amplitude thresholding alone did not suffice to produce segmentation as defined above (Pavlidis, 1980). Thresholding merely separated individual pixels by brightness level, regardless of connectivity between the points. The connected-component labeling algorithm may be used in conjunction with thresholding to satisfy the connectedness condition above. This was different from what had been used in prior research.

In local thresholding, Chow & Kaneko (1972) broke down the original image into smaller sub-regions, and gave each a threshold individually. In many cases, this caused discontinuities at the borders between sub-regions. To eliminate this problem, Chow and Kaneko interpolated the threshold values from neighboring sub-regions. The researchers divided the image into smaller regions with a 50% overlap (Chow & Kaneko, 1972). Chow and Kaneko computed the grey-level histogram for each region, and those with large variances were selected. For all other regions, the researchers interpolated the threshold from adjacent regions of a known threshold. Then, the researchers estimated the component distributions and the mixture of the probability density functions for each selected region using curve fitting. The resultant mixtures of estimated distributions were tested for bimodality. The method of maximum likelihood calculated the threshold from the estimated distributions for every histogram with appreciable bimodality. Chow and Kaneko demonstrated their method on cineangiogram images of the left ventricle of the human heart. The results successfully demonstrated the feasibility of the method for low-quality images.

Pappas (1992) presented a local thresholding method that was a generalization of the K-means clustering algorithm presented previously (as cited in Tou & Gonzalez, 1974). Unlike the K-means, Pappas' method was adaptive, and included spatial constraints by the use of a Tou and Gonzalez (1974) random field model. The intensity functions were initially constant in each region, and equal to the K-means cluster centers. In each iteration, Pappas updated the intensities by averaging over the region, and progressively decreasing the size of the overall region (as cited in Tou & Gonzalez, 1974). The algorithm ended when Pappas reached the minimum desired size.

Histogram Analysis

Although the first-order statistical functions described above provided a simple

solution, the predictions varied over the many regions of an image (Pavlidis, 1982). A

histogram of a digital image provided more information with minimal computational

complexity (only one pass of the image). It was a discrete function whose x-axis

represented discrete brightness levels and whose y-axis represented the number of pixels

with brightness value. The histogram peaks identified ranges of brightness corresponding

to the different objects or regions. Threshold values therefore may be selected according

to the locations of the valleys in the histogram, which result in a more sensitive but still

restricted segmentation (Pal & Pal, 1993).

The K-means clustering algorithm described below approximated the optimal peaks in

an image histogram. Each cluster represented a group of pixels with homogeneous grey-

levels. K was the variable number of clusters desired. If one wanted the region to be

split into two grey levels for example, K would be equal to 2. The algorithm goes as

follows (Tou & Gonzalez, 1974):

1. Take the first K pixels in the region. Assign them to be the initial cluster

   centers. $z_1(1), z_2(1), \ldots, z_k(1)$. Set iteration number k=1.

2. kth iteration: distribute the pixels among the K cluster domains $S_j(k)$ using

$$x \in S_j(k) \ if \ \left\| x - z_j(k) \right\| < \left\| x - z_i(k) \right\|, \ \forall i=1,2,\ldots,k, i \neq j$$

3. Update cluster centers $z_j(k+1)$ with N_j number of samples in domain $S_j(k+1)$:

$$z_j(k+1) = \frac{1}{N_j} \sum_{x \in S_j(k)} x, \quad j=1,2,\ldots,k$$

4. If cluster centers $z_j(k+1)$ differ from the previous iteration, $z_j(k)$, for j then step

2.

The resulting K cluster centers represented the most populated brightness; also the K was the highest peaks in the histogram. The number N_j of pixels in each cluster may be considered the "weight" of the cluster brightness. Evaluating a weighted average such as where the Z_j are ordered

$$x_j = \frac{N_j z_j(k) + N_{j+1} z_{j+1}(k)}{N_j + N_{j+1}}, \quad j = 1, 2, ..., K-1$$

of these brightness levels, one could consider this value to correspond to a valley floor in the histogram. For K=2, the clustering algorithm returned to peak brightness, of which the weighted average was assumed to be the floor of the valley brightness (Tou & Gonzalez, 1974). Thresholding the image with this valley floor value generates input for the next section's algorithm, connected-component labeling. As an example to demonstrate thresholding, Figure 7 presents the results of manual thresholding an image with two objects of different brightness.

Boundary-Based Methods

Unlike the previous two methods, boundary-based sought the discontinuity of brightness levels rather than their homogeneity. Although point, line, and edge discontinuities could be defined, the most predominant method was edge detection, because real images are not composed of simple points and lines (Li & Zhang, 2006). Applying edge detection operators on an image resulted in a characterization of object borders, and thus a segmentation of the image (Gonzalez & Woods, 2002).

Edges were points of abrupt change in brightness levels in digital images (Jain, 2002), and this led to the definition of a differential operator on digital images. Gonzalez and Woods (2002) defined the *gradient of an image* f(x,y) at location (x,y) as the image

vector, which points in the direction of the maximum rate of change of f(x,y) with respect to x,y. The magnitude of ∇ f equaled:

$$\nabla f = \sqrt{G_y^2 + G_x^2}$$

which measured the rate of increase of f(x,y) per unit distance in the direction of ∇ f. The Sobel operator did well in approximating the value of the gradient, and simultaneously had a smoothing effect that diminished the noise enhanced by the derivatives. The digital derivatives were based on the Sobel operators defined by Gonzalez and Woods (2002).

It was Pavlidis' (1980) opinion that merely detecting areas of high contrast in an image did not necessarily result in a complete segmentation (Pal & Pal, 1993). The windows in the shade in the middle-left of Figure 7(a) serve as a prime example of this case; they were accurately defined in the original, but were not completely surrounded by one true, single edge in the gradient image (see Figure 7(b)). Jain (2002) also stated that since these local operators were applied to a very limited region, they produced false contours. Contours that corresponded to object boundaries and contours that result because of noise within the small operator mask region could not be distinguished.

(a)   (b)

Figure 7. Edge detection example (Kunt et al., 1985), (a) Original picture of an M.I.T. Building, and (b) After applying the Sobel operators.

Contour Following

Pavlidis (1980) developed an algorithm to extract the contour of a single region

starting at the upper-left member pixel. A "walk" is taken clockwise around the region,

and the algorithm stops when the initial coordinates are met again (Pavlidis, 1980). The

single pixel region triggers the "insanity" (circling) condition, and may have been

encoded simply with a direction operator pointing upward, since all chain-codes begin at

the uppermost point of a region (Pavlidis, 1980). A sample run would look like the

image shown in Figure 8. Figure 9 presents the contour tracking algorithm.

(a)   (b)   (c)

Figure 8. Contour tracing graph (a) Region (shaded) and background; (b) result of
contour tracing with absolute direction operators; and (c) contour tracing with relative
direction operators (the algorithm for contour tracing is presented in Figure 9).

```
x=x_orig, y=y_orig  //set initial point to (x_orig,y_orig) upper-leftmost point
first=true              //flag indicating first iteration
dir=oldir=0;            //i.e. start off going in an easterly direction
                        //(all directions are assumed modulo 8 see Figure 8)
while((x!=x_orig)or(y!=y_orig)or(first==true){
        found=false;
        dir=dir − 2;  //turn dir anticlockwise twice
        while(found=false and insanity<8){
                insanity++
                set(newx, newy)to the coordinate of the dir-neighbor of (x,y)
                if ((newx,newy) is a member or the region){
                        reldir=dir-oldir;
                        export symbol [reldir] (relative direction operator)
                        oldir=dir;
                        (x,y)=(newx, newy);
                        found=true;
                }
                else
                        ++dir;  //turn dir clockwise
```

```
}
if(insanity==8)   //single-pixel regional growing around in circles is insanity!
        export symbol [-1]
         // any dir-op going up on the first symbol is an exception
      first=false;
}
```

Figure 9. Contour tracing algorithm.

Region-Based Methods

Pixel aggregation techniques started with one point, and grew in all directions around

that point to form a region, giving a seed point to neighboring pixels (Saidin et al., 2010).

Those pixels that succeeded became part of the same region as the original pixel. A

growth acceptance criterion compared the candidate pixel's brightness with the region

average. For example, Kroon (2008) applied this procedure to the neighbors of the

newly-grown region, until the growth acceptance criterion was not being met by any

more neighbors. Then, after Kroon found a complete region, the researcher selected

another seed point, disjoint from all previous regions, and repeated the process. Figure
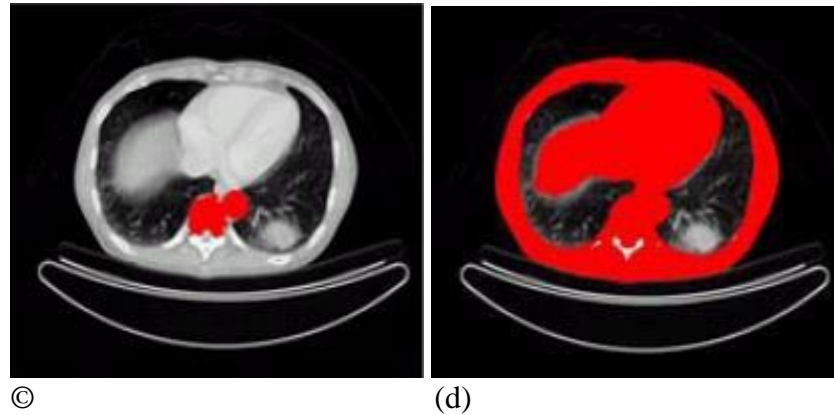
10 shows how the algorithm progresses.

(a)   (b)

©   (d)
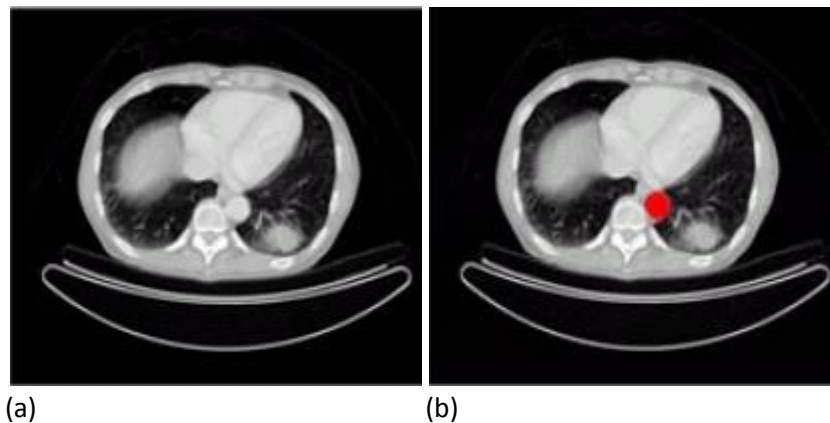
Figure 10. Regional growing (Kroon, 2008); (a) center of the lung exist a seed point; (b) region has grown in all directions to form a round shape; (c) the brightness in the center was not accepted by the uniformity predicate; and (d) final stage where no further growth is permitted by the uniformity predicate.

Regional-based methods use a neighborhood to threshold each pixel individually (Kroon, 2008). The neighborhood was centered on each pixel, and a global thresholding technique found a threshold value for the pixels within the neighborhood. Kroon (2008) used this threshold value to threshold the center pixel, and then the center of the neighborhood was moved to an adjacent pixel. Thus the threshold for each pixel was based on local properties. The method was a local thresholding technique that used a global technique to threshold each local region (Kroon, 2008).

Assuming an N x M image matrix X with a brightness function f(x, y), a pixel P at coordinates (x, y) has horizontal and vertical neighbors with coordinates:

$$(x-1, y), (x, y-1), (x, y+1), (x+1, y)$$

and has four more diagonal neighbors with the coordinates:

$$(x-1, y-1), (x+1, y-1), (x-1, y+1), (x+1, y+1)$$

A pixel Q is a 4-neighbor of pixel P if its coordinates match any of the horizontal or vertical neighbors of P. A pixel Q is an 8-neighbor of pixel P if its coordinates are among either the connected horizontal, vertical, or diagonal neighbors of P (see Figure
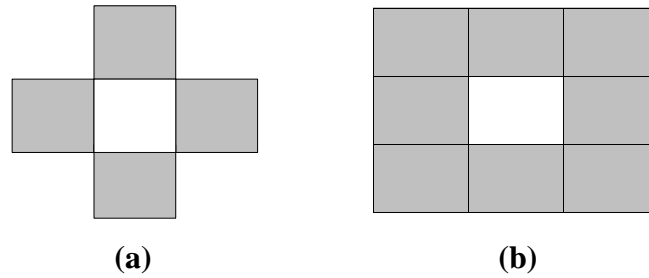
11).

         (a)   **(b)**

Figure 11. 4- and 8-Neighbor Connectedness (Jain, 2002); (a) A central pixel and its 4-connected neighbors, and (b) a pixel and its 8-connected neighbors.

A region is a set of pixels. A pixel X_i in region R z-connects to pixel X_j in R if a path exists in R such that all pairs of successive path pixels X_k and X_{k+1} are z-neighbors. A region R is a connected region if the set of pixels in R had the property that every pair of pixels is connected (Ballard & Brown, 1982).

Regional growing of neighboring pixels of similar values can be found in connected-component labeling (Ballard & Brown, 1982). According to Ballard and Brown (1982), all pixels that had at least one neighbor defined the contour or perimeter of an object. This was not in all the pixels that lay on the image frame. The contour of an object describes the shape of the object. Since, in the quadtree example, the researchers processed first the predominant brightness of the entire image, followed by the predominant brightness of the four major quadrants, followed by the predominant brightness of the 16 sub-quadrants and so on, the effect was one of a gradually resolving image with finer and finer detail at each step (Ballard & Brown, 1982). This allowed a compressor, and then de-compressor, to process an image only to the detail they were required to.

Region-growing techniques did not place a restriction to the shape or size of an image

(Nock & Nielsen, 2004). Most split-and-merge algorithms employed tree structures with restrictions to the shape that a region could take (e.g., square). This meant that in order for tree-based techniques to be efficient, the objects' contours had to be either horizontally or vertically oriented, and had to be in a position corresponding to the borders of the tree nodes (Wu, 1992).

Researchers who used region-growing implementation could have implemented extensive memory that was required to hold states awaiting expansion, in addition to those that already had expanded (Lin, Jin, & Talbot, 2001). Apart from being time consuming, an unpredictable number of regions generated from regional growing. Lin, Jin, and Talbot (2001) could have utilized an adaptive algorithm to control these features of the output, but it came with a high cost because of its high computational complexities. One advantage that became more evident was that region-growing techniques did not pose any restrictions on maximum region size or shape (Lin et al., 2001).

In region splitting and merging, the algorithms took a different approach to subdividing an image by utilizing a tree structure [see Figure 12(a)]. The entire image was represented by a root node R that was then subdivided recursively into a fixed number of siblings. The individual pixels were the leaves of the tree. This formed a quadtree structure as depicted in Figure 12(b). Having a fixed number of siblings per node, and preset shapes for each level of the tree, made retrieval of region data simple and straightforward. The following is an example of a standard splitting technique for that representation of a quadtree from Gonzalez and Woods (2002):

1. Let R represent the entire image region and select the predicate P. R is to subdivide successfully into smaller and smaller quadrant regions so that, for

any region R_i, $P(R_i)$= TRUE.

2. If $P(R)$ = FALSE, divide the image into quadrants.

3. If $P(R)$ is FALSE for any quadrants, apply subdivision method recursively on that sub-quadrant. It should be noted that by definition, it is not possible that $P(R)$ = TRUE and $P(R')$ = FALSE for some child (sub-quadrant) R' of R.

4. Stop when no further splitting is possible.

For example, the quadtree structure divided a parent square region into four square sub-regions, the quadtree structure divided into two geometrically equal rectangles (alternating orientation), etc. Each node of the quadtree had four children. The root node R represented the entire image; its children represented the four quadrants of the entire image; their children represented the 16 sub-quadrants; the children of those represented the 64 sub-sub-quadrants, and so on. If a non-leaf node had its corresponding image data area conforming to a uniformity predicate, then the average brightness was stored in the parent and the children were deleted. Thus, if an entire quadrant (sub-quadrant, sub-sub-quadrant, etc.) of the image was a single brightness (e.g., white), that information could be seen as a single quadtree node.
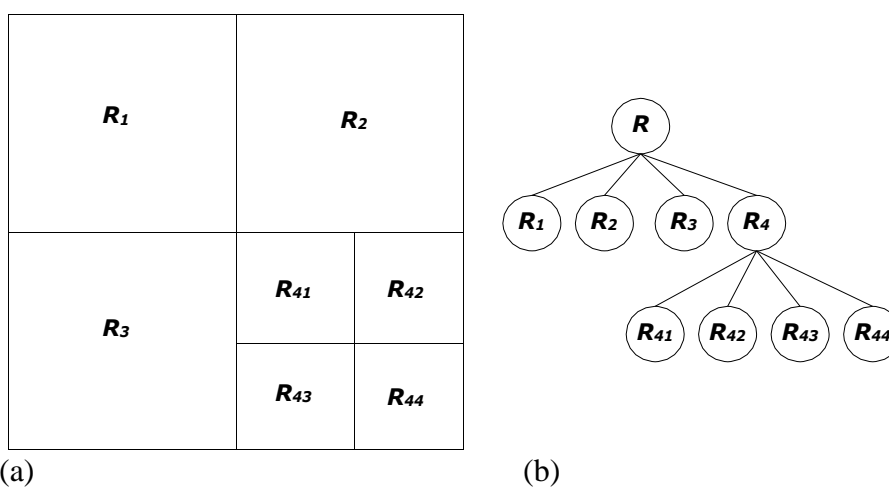

(a)   (b)

Figure 12. Partitioned quadtree image (Gonzalez & Woods, 2002); (a) Partitioned image,

and (b) corresponding quadtree.

Splitting schemes started with evaluating the root node for uniformity, and recursively split nodes for which the predicate was not satisfied (Gonzalez & Woods, 2002). The split was a division into a number of sub-regions specified by the technique being used. For example, in the quadtree technique Gonzalez and Woods (2002) divided a square region (required) by splitting the region horizontally and vertically at the middle of each side to form four square sub-regions. The researchers also checked each sub-region for uniformity, and if it was non-uniform, its sub-regions were generated recursively (Gonzalez & Woods, 2002). This process repeated until all regions satisfied the predicate P for uniformity criterion. Splitting an image implied starting at the top of the tree and traversing downward. On the other hand, the merging approach started at the leaf nodes of the quadtree when all leaf nodes satisfied the uniformity criterion i.e., TRUE under P, and merged upward in the hierarchy of regions. The following was a standard merging predicate from the work of Gonzalez and Woods (2002).

The researchers merged only adjacent regions with combined pixels that satisfied predicate P two adjacent R_j and R_k were merged only if $P(R_j \bigcup R_k) = TRUE$ (Gonzalez & Woods, 2002).

1. Split into four disjoint quadrants any region R_i for which $P(R_i) = FALSE$.

2. When no further splitting was possible, merged any adjacent regions R_j and R_k for which $P(R_j \bigcup R_k) = TRUE$.

3. Stopped when no other merging was possible.

Gonzalez and Woods (2002) defined the following conditions to merging where $P(R_i)$ was a logical predicate defined over the points in set R_i and $\varnothing$ was the null set.

1. $\displaystyle\bigcup_{i=1}^{R} R_i = R.$

   This first condition indicated that the segmentation must be completed; that was every pixel must have been in a region.

2. R_i was a connected region i = 1, 2,…, n.

   This condition required that points in a region were connected in some predefined sense (e.g., 4- or 8-connected).

3. $R_i \cap R_j = \varnothing$ for all i and j, i $\neq$ j.

   The third condition indicated that the regions must be disjoint.

4. $P(R_i)$= TRUE for i = 1, 2,…, n.

   The properties that must be satisfied by the pixels in a segmented region for example $P(R_i)$ = TRUE if all pixels in R_i had the same grey level.

5. $P(R_i \cup R_j)$ = FALSE for any adjacent region R_i and R_j.

   Finally, condition five indicated that adjacent regions R_i and R_j were different in the sense of predicate P.

Instead of starting at the top (whole image) or the bottom (pixel level) of the tree, starting at an intermediate level and thus applying a split-and-merge algorithm was more advantageous (Gonzalez & Woods, 2002). This algorithm traversed the nodes at the preset intermediate level, and either split or merged according to the degree of uniformity of the regions (nodes). If a region had a variance that was higher than a certain threshold, that region was split. If Gonzalez and Woods (2002) found two regions to have highly-similar features (i.e., a merge test passed the predicate), the regions were merged. The split-and-merge techniques were superior; researchers found a suboptimal solution tree

quicker (Pavlidis, 1980).

Single- and multiple-amplitude thresholding techniques generated thresholds based upon the statistical features of a local region. They could be implemented so that a new threshold was conceived for a new sub-region. A method for deriving a variance threshold, however, was quite different. Splitting regions into the least number of sub-regions was preferable to splitting to produce a large number of sub-regions. This was because a smaller number of sub-regions were then split even further to produce the same sub-regions that were originally conceived.

In split-and-merge, regions depended on some uniformity criterion. Uniformity criterion is a subset property, such that a uniformity predicate exists if $P(R_1)$ and $R_2 \subset R_1$, then $P(R_2)$. This is usually aimed at finding some uniform intensity, or uniform-like background image regions. Pavlidis (1982) investigated variance thresholding techniques as methods for defining suitable regions and sub-regions for splitting algorithms. In split-and-merge schemes both methods had advantages and disadvantages. Merge-only algorithms must exhaustively test uniformity criteria on single pixels, whereas split-only algorithms created the need to recheck the resulting sub-blocks for uniformity. Both methods required more time to reach a final result because of the expensive nature of their operations (Pavlidis, 1982).

Connected-Component Labeling

The amplitude-splitting algorithms may have resulted in the same labels being applied to regions that were not necessarily connected. Figure 13I contains two disjoint regions that had been labeled by unity. The same two objects in Figure 13(b) had been assigned different labels, namely 1 and 2. The latter complied with the definition of image

segmentation, but the former segmentation violated the connectedness rule.

Connected-Component Labeling with Absolute Threshold

The connected-component labeling algorithm assigned individual region numbers to connected groups of pixels, i.e., it takes Figure 13(a) as input and outputs Figure 13(c) (Pavlidis, 1980).
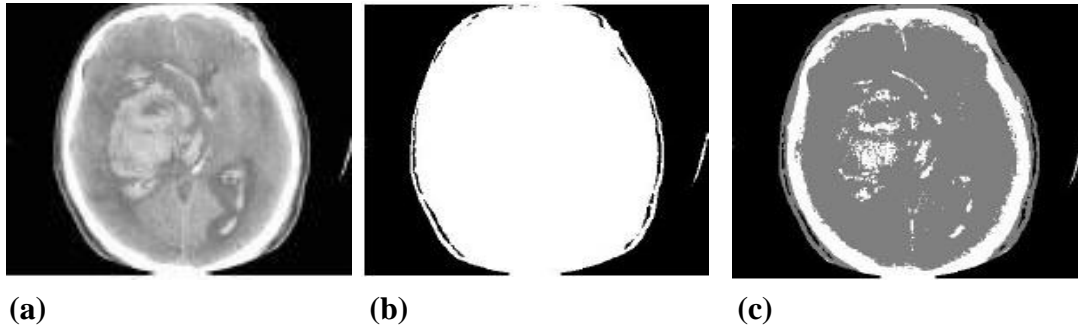

    (a)   **(b)**   **(c)**
Figure 13. Multilevel thresholding example (Csetverikov, 2012); (a) Original image, (b) Result of bi-level threshold, and (c) Result of a multilevel threshold.

Its sequential, top-left to bottom-right raster scan method was easy to implement. A connected-component labeling algorithm mapped a binary input image f(x,y) to a 4-connected labeled matrix s(x,y) (Ballard & Brown, 1982). A second pass was necessary because of V-shaped regions in the image. The algorithm was followed by a sort routine that identified unique labels for all the labels marked equivalent. Then, in a second pass of the matrix s, the equivalent labels were assigned to their corresponding unique label. The algorithm accepted binary input matrices, which meant this ideally could have been applied to the output from an amplitude thresholding process.

Connected-Component Labeling with Variance Threshold

The algorithm also could be extended to work on grey-level images by relaxing the definition of uniformity. Instead of comparing the upper and left neighbors' equality with the central pixel's brightness, the researcher of this study compared the difference of

their brightness with a threshold, basically using a local threshold as explained in the Amplitude Thresholding section.

The variance threshold allowed for some degree of "looseness" among neighboring pixels in this manner. Ballard and Brown (1982) considered the labeling to be a simple region-grower for non-binary images. The researcher of this study described characteristics of region growers in further detail earlier in the Region Based Method section. Both cases of connected-component labeling in this study were local algorithms because they involved only one pixel and its neighbors.

Classifiers

A computing task of much interest in the field of machine perception, both from a practical and theoretical standpoint, is that of pattern classification. The task of a pattern classifier is to make a decision about which class a given sample represents based on information extracted from that sample. These classifiers perform a transform, or mapping, from a set of feature values into one of a finite number of classes (White, 1990).

The proper construction of the training set is crucial to successful training. One of the criteria to be met for proper construction of a training set is that each of the classes has to be adequately represented (Vasantha & Bharathi, 2012). A class that is represented less often in the training data may not be learned as completely or correctly, impairing the network's discrimination ability (Vasantha & Bharathi, 2012). This is because of the implicit setting of a priori probabilities, which results from unequal sample sizes. The degree of impairment is a function of (among other factors) the relative number of samples of each class used for the training set (Vasantha & Bharathi, 2012).

As in the case of statistical classifiers, one of the criteria to be met for proper construction of a training set for a BPN was that the image classification had been adequately represented—enough samples and an approximately equal number of samples per image classification (White, 1990). The number of samples per classification was implicitly represented in a network as a priori probability (Arora & Suman, 2012). The more often a classification was represented in the training data, the more the weights were adjusted to favor that classification. An image classification that was represented less often in the training data was not learned as completely or correctly, and the network's discrimination ability could be impaired (Arora & Suman, 2012). If the condition of equal representation of each output classification was not met in the original data, the simple and often practiced resolution was to duplicate training data for the minority classification, or reduce the population of the over-represented classification until adequate representation of all classification was achieved (Arora & Suman, 2012).

Using this method, data were either duplicated or discarded to provide adequate minority image classification representation. In autonomous systems, the memory required while performing duplication in the training sets, or the intelligence to select vectors for removal, could be impractical to implement. Miyake et al. (2009) used image classifiers to find false positives and to detect statistical features based on shape of four-density features. When using decision tree classifiers to predict a particular output of class that is true, the decision tree classifiers will predict that all remaining output classes are false (Russell & Norvig, 2003). Olson and Delen (2008) found that for each classifier the following prediction would be used:

- True positive (tp) is when the classifier output correctly predicts that the input

is true.

- True negative (tn) is when the classifier output correctly predicts that the input is not true.

- False positive (fp) is when the classifier output incorrectly predicts that the input is equal to a true output.

- False negative (fn) is when the classifier output incorrectly predicts that the input is not equal to the true output.

Decision Tree

A *decision tree paradigm* provides a useful tool for solving certain types of classification problems. The overall hierarchy is a *tree*, composed of nodes and branches. Internal nodes corresponded to decision locations in the tree. *Branches* connect each parent node to its child nodes (or children) when moving from top to bottom, and connect child nodes to parent nodes when moving from bottom to top. The node at the top of the tree is referred to as the *root node*. Nodes with no children are called *leaves;* these also are referred to as *terminal nodes* and hold all of the possible solutions that can be derived from the tree. With the exception of neural networks, decision trees are the most widely utilized tool within data mining applications, and are a useful technique for image classification (Russell & Norvig, 2003).

The advantage of decision tree methods over other modeling tools, such as neural networks, is that it produces a model that may represent English rules or logic statements. Another advantage associated with decision trees is that no assumptions about the data are necessary. Decision trees can be used to automatically classify the input variables (test sets) based on the overall strength of their influence on the decision tree (training

set). Figure 14 presents the algorithm for decision tree learning.

```
function Decision-Tree-Learning (examples, attribs, default) returns a decision tree
    inputs: examples, set of examples
            attribs, set of attributes
            default, default value for the goal predicate

    if examples is empty then return default
    else if all examples have the same classification then return the classification
    else if attribs is empty then return Majority-Value (examples)
    else
        best ← Chose-Attribute(attribs, examples)
        tree ← a new decision tree with root test best
        m ← Majority-Value(examplesᵢ)
        for each value vᵢ, of best do
            examples ← {elements of examples with best =vᵢ}
            subtree ← Decision-Tree-Learning (examplesᵢ, attribs – best, m)
            add a branch to tree with label vᵢ and subtree subtree
            return tree
```

Figure 14. Decision tree learning algorithm (Russell & Norvig, 2003).

Nonetheless, one can certainly use rule-based methods with a classifier based on an

artificial neural network (ANN) to find candidates for cancer. This technique can provide

a useful computer-aided detection (CAD) application for feature detection of malignant

growth to help radiologists and physicians detect and diagnose possible cancerous

growths. Gilbert, Astley, and Gillian (2008) used CAD for screening mammography

images. They reported that CAD had potential cost savings, and that using two kinds of

readers—human and machine—had the potential to improve cancer detection rates.

ANNs comprise a class of computing paradigms that have recently spawned a great deal

of research. One widely used neural network paradigm is the multilayer perception,

which employs back-propagation of errors learning, or Back-Propagation Network

(BPN). BPNs have proven very useful in the field of pattern image classification by

mapping image input patterns into one of several categories. Rather than requiring

specific programming, BPNs learn new mappings by exposure to a training set, which comprises a collection of input image pattern samples matching with their corresponding output images.

Ultimately, the present study used decision trees in lieu of an ANN. Gupta, D. L., Malviya, A. K., & Gillian, Singh (2012) examined J48 methods of classification and observed that the J48 methodology had maximum accuracy and minimum error rate. On the basis of accuracy measures of the classifiers, one can provide guidelines regarding fault-prone prediction issues of any given data set in the respective situations.

The Miyake et al. (2009) paper discussed using rule-based methods and artificial neural network classifiers to detect features based on shape and density in order to find lung nodules. In contrast to ANN, decision trees have become the *de facto* standard when detecting features based on shape and density in WEKA 3.6.10 (Arora, R., & Suman, 2012). This paper illustrates how neural networks give a lower classification error rate than decision trees but require longer learning time. Additionally, a WEKA 3.6.10 decision tree can convert into a set of (mutually exclusive) rules, with each rule corresponding to a tree branch.

Abeer Y. Al-Hyari, & Ahmad M. Al-Taee (2013) used the open source WEKA 3.6.10 software in their study for performance comparison and evaluation purposes. The obtained results showed that the developed decision tree algorithm represents the most accurate Chronic Renal Failure (CRF) classifier (92.2%) when compared to all other algorithms/implementations involved in this study.

Chapter 3

Methodology

General Approach

The goal of this approach was to help radiologists detect cancerous tissue in CT and MRI mammography images with CAD system design. Doctors order a patient to take a CT or MRI mammogram every year for routine checkups, and radiologists routinely screen hundreds of these mammogram images for cancerous tissues. The proposed methodology could help radiologists detect cancerous tissues by using a CAD system to help locate the region of interest. The experimental design is explained in detail at the beginning of Chapter 4.

This methodology could improve breast cancer detection by automatic detection of cancerous tissue from CT and MRI mammography images by enhancing the CT and MRI mammography image using image registration and subtraction. Next, in this dissertation the researcher used research thresholding techniques, such as regional growing, to find the region of interest. Finally, the researcher employed decision trees classifications to reduce false positives and false negatives. Thus, this dissertation is a three-stage approach to detect malignant growth in CT and MRI mammography images (see Figure 15).
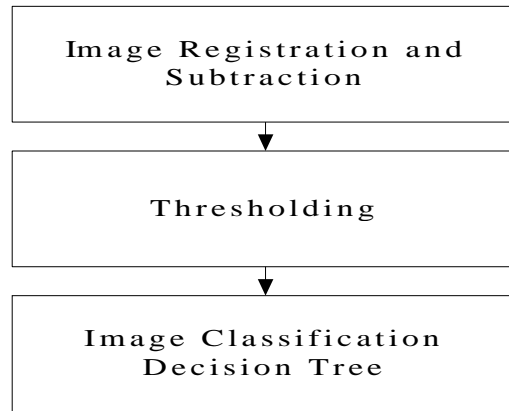
```
┌─────────────────────────────┐
│  Image Registration and     │
│       Subtraction           │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│        Thresholding         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│   Image Classification      │
│      Decision Tree          │
└─────────────────────────────┘
```

Figure 15. Methodology overview.

Miyake et al. (2009) used the same type of methodology in rule-based, artificial neural network classifiers to detect features based on shape and density and to find lung nodules. This dissertation involved the same type of CAD methodology to detect cancerous tissue in CT mammography images, and also examined MRI images. The researcher also used two types of subtraction methods and different thresholding methods.

The work of Miyake et al. (2009) was similar to this study, but differs technically in that the researchers considered registration with CT images, whereas this work involves with both CT and MRI images. In addition, Miyake et al. only used temporal subtraction. This dissertation considered both temporal subtraction and dual-energy subtraction techniques with mammogram images taken some time apart. By using registration and subtraction, Linh and Linh. (2010) stated that one of the keys of usefulness of subtraction was the enhancement of differences between images.

Miyake et al. (2009) only used histograms for thresholding. This dissertation considered both global and local thresholding algorithms. The thresholding methods proposed have no restrictions on region shape or size, and are tailored to allow constraints to be applied and approximated with ease. The local region-growing

techniques combine the flexibility of quadtree-based algorithms with the non-shaped-restricted features of region-growing schemes. Regional splitting algorithms, such as the one proposed, are generally noise-tolerant, and thus less prone to information loss from image pre-processing (noise filtering).

Image Registration and Subtraction

To achieve a high degree of accuracy in the registration, researchers used external markers (e.g., stereotactic frame systems) for a long time—since 1947. The idea of using marker implantation in order to provide for image registration was simple, and had some advantage over techniques that rely on natural landmarks. For example, the marker could be tailored to serve the purpose of accurate image registration; once the appropriate marker had been devised it was the same for every patient.

In the last decade, techniques for retrospective registration of tomographic images began to appear. The retrospective techniques mainly rely on the localization of anatomical features and provide a non-invasive method of registration (Linh & Linh, 2010). Linh and Linh (2010) used the registration tool from Matlab for image registration by utilizing multimodal image registration to optimize the alignment of CT and MRI data. Figure 16 shows the Matlab tool with image registration of an MRI brain scan. This tool allowed alignment of 3D images from different medical imaging modalities. In this dissertation, the researcher used the same tool for registration of CT and MRI mammographic images.
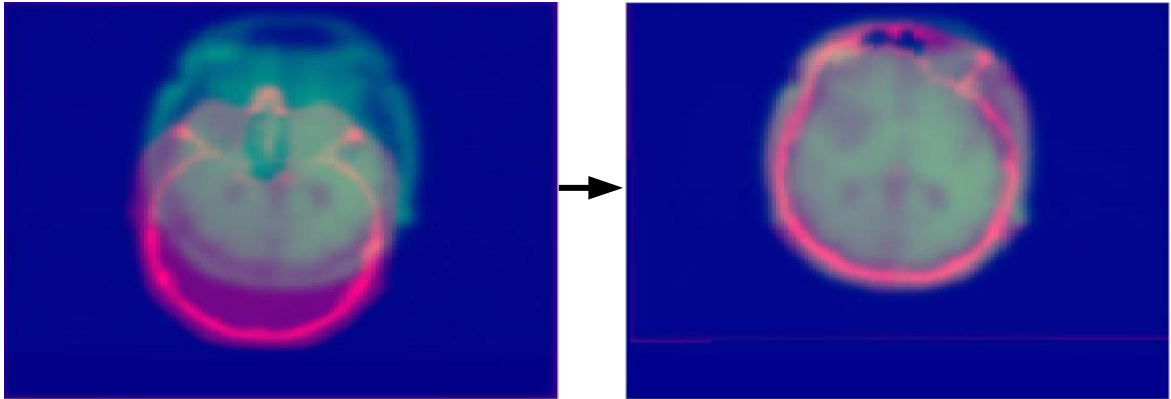
Figure 16. Image registration with Matlab.

The second part of image enhancement requires image subtraction. Linh and Linh (2010) described *image subtraction* as being an enhancement process that subtracts the difference to bring out greater detail in the image. The image subtraction method removes the local change in illumination over a gray-level image. The researcher of this study used two types of subtraction techniques for CT and MRI images: temporal subtraction and dual-energy subtraction. The dissertation researcher investigated both techniques. The researcher used temporal subtraction experiments with Matlab, and the dual-energy subtraction was acquired from image databases and compared with temporal subtracted images.

By using low-pass spatial filtering, the researcher created a new image with the overall contrast and average intensity of the original image. By subtracting it from the original image, the regional variations within the illumination were removed and the resulting image was then thresholded using a global technique (see Figure 17). Miyake et al. (2009) used temporal subtraction images obtained by the voxel-matching technique and threshold techniques to find lung nodules.
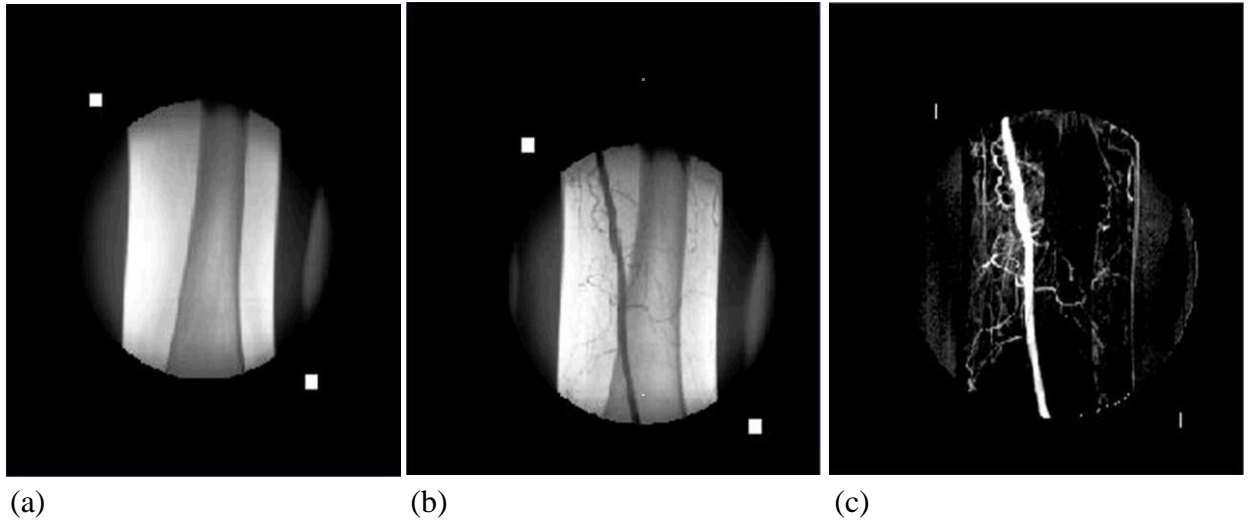
(a) (b) (c)

Figure 17. Subtraction images with Matlab (Linh, 2010); (a) Mask image, (b) Image taken after contrast, and (c) Image with mask subtracted out.

Methods for Identifying ROI

The work of Tourassi, Vargas-Voracek, and Floyd (2003) highlights the methods used to identify ROI. In their work, they developed content-based image retrieval (CBIR) using DDSM with over 1,009 mammographic images from the database to find ROI. Vasantha and Bharathi (2011) stated DDSM contained more than 2,620 cases acquired from Massachusetts General Hospital, Wake Forest University, and Washington University in St. Louis School of Medicine, and that the DDSM comprised patients of different ethnic backgrounds. Eltonsy, Tourassi, and Elmaghraby (2007) presented a technique for automatic detection in mammographic masses in images that were cropped into regions containing malignant growth. These researchers extracted morphological characteristics from these regions to find ROI, and they then used a minimum distance classifier in 540 images containing malignant masses, 270 images containing benign masses, and 164 normal images from the DDSM database (Eltonsy, Tourassi, & Elmaghraby, 2007).

In this dissertation, the approach for finding the ROI is a five step process. First, one must apply the threshold method to the subtraction image. From this thresholded image, the ROI gets identified, and subsequently registered, with the original image. Next, feature vectors from the ROI (relative to the original image) get extracted from the non-ROI. The final step involves classifying the feature vectors using a decision tree.

For all images in this study, the researcher completed Classification Analysis, checking for accuracy of segmentation and quality of classification. Additionally, the researcher computed twenty-five texture feature type values for the *malignant* and *benign* areas of each image. This combined the test set and training set from thresholding to produce the precision, recall, and f-measure. An image classification decision tree calculated the final values from WEKA 3.6.10 along with WEKA J48 classifier and cross-validation 10 fold on all experiments.

In this dissertation, the Image Processing Toolbox supported different types of image file formats for medical images, such as tomographic images (Matlab, 2012). The Image Processing Toolbox also provided comprehensive image display capabilities such as zoom, pan, and examination of regions of pixels. The graphic user interface allows interactively identifying the Region of Interest (ROI), adjusting contrast, cropping, and measuring distance (Matlab, 2012).

For deriving the training sets, the researcher examined possible classifications of the feature vectors of an image based on the ROI in the image (Vasantha & Bharathi, 2012). The original images have an ROI of the malignant region that a radiologist previously identified, and this ROI then carried over to the thresholded images. The extraction of

the ROI was necessary to analyze the feature vector, and to find the prominent features

that were representative of the classes of the images.

Thresholding

There were eight types of thresholding methods in this dissertation: amplitude

thresholding; global thresholding; local thresholding; connected-component labeling with

absolute threshold and variance threshold; threshold in histogram analysis; thresholding

with K-means clustering; regional growing thresholding algorithm; and split-and-merge

thresholding algorithm functions based upon the density of brightness values present in a

region. After considering the results (shown in the results section Chapter 4 Results), the

researcher chose four thresholding methods for further experimentation: amplitude

thresholding, local thresholding, K-means clustering, and regional growing thresholding.

Out of the original eight thresholding methods, the researcher deemed these four

techniques acceptable for the next stage of image classification.

To clarify, the researcher excluded the global contour method as it did not detect one

of the cancerous tissues shown. Notably, it did show the full contour of the image and

followed the Otsu method. Nevertheless, the global thresholding, connected-component

labeling, and split-and-merge methods will not appear in the Appendix since all results

were the same due to all of the images having the same brightness level in the foreground

or background. Additionally, these different thresholding methods were tested and found

to be unusable for each image type.

One should also note that threshold based techniques oftentimes get referred to as

histogram-based methods, which make decisions on local pixel information and are

effective only when the intensity levels are far outside the range of the levels in the

background. Also of importance to note, when using amplitude thresholding, the most noticeable difference between temporal subtraction images and dual-energy images is the foreground truth, in which temporal subtraction is white and the dual-energy subtraction is black.

The second stage examined the four different thresholding techniques. The amplitude thresholding method manipulated objects and backgrounds in such a way that object and background pixels had grey levels grouped into two dominant and different modes. In these cases, the researcher could extract the objects from the background using a threshold that separates the modes. The local thresholding introduced posed no restrictions on region shape or size, due to the fact that it maximized edge features by thresholding local regions separately. Regional-growing methods used recursive splitting to partition the image top-down by using the average brightness of a region. Importantly, the researcher applied each thresholding method to each of the three image types. A more detailed description of each of these methods is contained in the following paragraphs.

Amplitude Thresholding

To reiterate, the amplitude thresholding method manipulates objects and backgrounds in such a way that object and background pixels contain grey levels grouped into two dominant and different modes. In these cases, the researcher extracted the objects from the background using a threshold that separates the modes. In other words, amplitude thresholding segments an image by setting all pixels whose intensity values are above a threshold to a foreground value, and all the remaining pixels to a background value. As such, amplitude thresholding typically takes a grayscale image as input and, in the

simplest implementation, outputs a binary image representing the segmentation. For each

segment, each pixel is classified relative to (above or below) the threshold to be

calculated. If the pixel value is below the threshold, it is set to the background value,

otherwise it assumes the foreground value. Simple and fast functions include the *mean* of

the *local* intensity distribution.

When implemented, an amplitude thresholding algorithm separates the foreground

from the background with non-uniform illumination. Importantly, illumination comes

into play in the mammography images after thresholding. Illumination can exist in any

image and represents one of the key elements in creating an image, particularly based on

the interplay of shadow and high intensity light. As such, one technique for image

segmentation is amplitude thresholding (Gonzalez et al., 2002). In some cases, an image,

f(x,y), was composed of objects and backgrounds in such a way that object and

background pixels had grey levels grouped into two dominant and different modes. In

these cases it was possible to extract the objects from the background using amplitude

thresholding that separates these modes.

Image thresholding operations followed the general relation for a threshold T

(Gonzalez et al., 2002). Point's coordinates, p(x,y) is a local function in the

neighborhood of pixel (x,y) (e.g., the average of the eight neighboring brightest pixels),

and f(x,y) is the grey level of the current point or pixel. T is equal to f(x,y), which is grey

level of (x,y). The application of T—classification of (x,y) with respect to T—depends

solely on f(x,y). In the case of 'local threshold', application of T depends on p(x,y), a

function of the neighborhood of (x,y). If T depends on p(x,y) in addition to f(x,y), it is a

local threshold. Additionally, T is said to be dynamic if it depends on all four of the components (Gonzalez et al., 2002).

Local Thresholding

The local thresholding method poses no restrictions on region shape or size, due to the fact that it maximizes edge features by thresholding local regions separately. In local thresholding the original image is decomposed into smaller sub-regions, and each is given a threshold individually. A pixel represents the center of a neighborhood if it achieves a local maximum (in brightness) relative to its neighbors. Furthermore, the NHOOD function determines which pixels are neighbors. In fact, NHOOD neighbors is a variable that provides a generalization of '8 neighbors' in that it defines the shape of each pixel's neighborhood. More specifically, NHOOD comprises an array of zeros and ones in which the nonzero elements specify the neighbors used in the computation of the local mean and standard deviation; as such, NHOOD defines the size of the neighborhoods. The size of NHOOD must be odd in both dimensions because it is centered on (x,y). In many cases, this caused discontinuities at the borders between sub-regions. To eliminate this problem, the researcher interpolated the threshold values from neighboring sub-regions. The approach used was the same one proposed by Chow and Kaneko (1972).

 K-means Clustering

K-means is a popular clustering algorithm for large datasets. This algorithm iteratively computes a set of k centers that implicitly represents a partition. Given any set Z of centers, each $z \in Z$ has a neighborhood defined as the set of data points that are closer to z than to any other center in Z. K-means starts with a set Z of centers and computes

their neighborhoods. In successive iterations, every center moves to the centroid of its neighborhood, and then one must re-compute the neighborhoods based on the updated positions of the k centers. This process continues until it satisfies a convergence criterion, such as when two successive iterations produce no changes to any of the k neighborhoods.

The collection of neighborhoods that results is taken to be the partition of the data points produced by k-means applied to the initial set of centers. Since some of the neighborhoods may be empty, the resulting partition comprises up to k nonempty clusters. Having applied k-means, pixels must now get classified into k classes. The graphic user interface was also used in K-means clustering as this allowed interactively looking at Region of Interest (ROI), adjusting contrast, cropping, and measuring distance (Matlab, 2012). Matlab used histogram-based methods, which make decisions on local pixel information and are effective only when the intensity levels are far outside the range of the levels in the background.

Interestingly, the K-means clustering algorithm approximated the optimal peaks in an image histogram. Each cluster represented a group of pixels with homogeneous grey-levels. K was the variable number of clusters desired. For example, if one wanted the region split into two grey levels, K would equal 2. (See Histogram Analysis Chapter 2.) The researcher used Matlab's 2013 Distance tool in the Imtool that allows for measuring distance in images by using the tool function in the toolbar or by selecting Measure Distance.

Matlab had some examples for K-means clustering located in the file exchange. The researcher conducted an analysis for the accuracy of segmentation and quality of

classification on this experiment. It often worked better than Otsu's method, which outputs larger or smaller thresholds on fluorescence images. While viewing the Dual-energy Subtraction Image, one can see that the last images showed the ROI. (See Appendixes A, B, and C for the greatest calculated precision, recall, and f-measure values.)

Region-Based Methods

The purpose of region growing is to partition an image into regions such that the pixels comprising each region satisfy a given similarity measure *P*. Specifically, a partition represents an equivalence relation over the pixels (every pixel belongs to exactly one region); every region is connected in a well-defined sense (generally 4- or 8-connected); every region satisfies the similarity measure *P;* and, generally, no proper superset of the region satisfies *P.* In contrast to split-and-merge segmentation, region growing proceeds bottom-up. To clarify, a selected seed pixel starts a region, and then adjacent pixels get added to the region as long as *P* is satisfied, and the process continues recursively on recently-added pixels until no further pixels can be added to the region without violating *P.* Additional regions successively grow from new (unvisited) seeds until a complete partition results.

The region-growing thresholding algorithm, obtained from the Gonzalez et al. (2012) paper, is an analysis for the accuracy of segmentation and quality of classification used on regional growing completed images. The Image Processing Toolbox also provided comprehensive image display for region growing capabilities such as zoom, pan, and examination of regions of pixels. Typically, one would let this image be an image where every pixel has its own label criterion. One uses the value of the similarity criterion to

decide between fusing segments and not fusing segments. In Figure 9 (Contour tracing algorithm) one can see how the region-growing algorithmic process functions, where P is coded. Pixel aggregation techniques started with one point, and grew in all directions around that point to form a region, giving a seed point to neighboring pixels (Saidin et al., 2010). The researcher would use similarity measurement using the Matlab GUI interface which allowed interactively looking at ROI, adjusting contrast, cropping, and measuring distance (Matlab, 2013).

When the ROI is identical to or similar enough to the one identified by the radiologist before thresholding, it is described as *holding the ROI*. This term identifies a region that has a malignancy, which refers to the area of the image that contains cancerous cells. If the ROI is very similar, but not identical to, the original ROI, this would be an area that holds the ROI because thresholding will change the look but not the overall size and degrees of the area. Matlab's 2013 Distance tool in the Imtool set allows for measuring distance in images by using the tool function in the toolbar or by selecting Measure Distance. This allows one to calculate and express in data units determined by the XData and YData properties, which by default is in pixels, the distance between two endpoints.

Image Classification

Decision tree learning consists of classifiers for instances represented as feature vectors (Gupta, Malviya, & Singh, 2012). In this approach, classification knowledge is first represented as a decision tree and then the tree is translated as a set of rules. The decision tree is constructed by sequentially selecting attributes based on an informational theoretical measure (Gupta et al., 2012).

Adequate representation of each of the image classifications is crucial to training. In

this study, an image classification that was represented less often in the training data might not have been learned as completely or correctly, impairing the network's discrimination ability. This stems from the implicit setting of a priori probabilities, which resulted from unequal sample sizes. The degree of impairment was a function of (among other factors) the relative number of samples of each image classification used for training.

A computing task of much interest in the field of machine perception, both from a practical and theoretical standpoint, is that of pattern classification (Gupta et al., 2012). The task of a pattern classification involves making a decision about which image classification a given sample represented, based on information extracted from that sample (Gupta et al., 2012). For the experiments for this methodology, the dissertation researcher used Digital Database for Screening Mammography (DDSM) images from the University of South Florida's mammography database. DDSM contained descriptions of breast lesions in terms of the American College of Radiology's (ACR) breast imaging lexicon called the Breast Imaging Reporting and Data System (BI-RADS) (Vasantha & Bharathi, 2011). These databases used the same image classification method that Miyake et al. (2009) used to find false positives and to detect statistical features based on shape and density.

For image classification, the researcher created J48 decision trees based on 21 features that were identified and obtained from the Matlab GLCM function, and four feature types that were obtained from the state of the histogram-based features. The J48 algorithm used a greedy approach in which the decision trees were constructed in a top-down recursive manner with a divide and conquer technique. Most algorithms for decision tree

methods take a top down approach. The algorithms start at the top of the tree with a training set of tuples and their associated class labels.

The researcher recursively divided the training set into smaller subsets as a tree was being built. The J48 algorithm consists of three parameters—attribute list, attribute selection method, and classification (Arora & Suman, 2012). The attribute list is a list of characteristics that describe the different tuples. The attribute selection method specifies a heuristic procedure for selecting the attribute that best discriminates the given tuples according to each class (Arora & Suman, 2012). The specified procedure employs an attribute selection method for selecting only those variables that contribute to information improvement that allow for multi-level splits. Each attribute selection method determines the splitting criteria. The splitting criteria tell which attribute to test at a node by determining the best way to separate or partition the tuples into individual classes (Arora & Suman, 2012). The researcher used WEKA 3.6.10 to create a decision tree training and test set using 25 features that were relevant at each node of the tree. WEKA 3.6.10 was used with the classifier J48 and cross-validation ten10 fold to create the final decision trees.

Vasantha and Bharathi (2011) used the Machine Learning package of WEKA to train their dataset, also using a decision-tree method. WEKA 3.6.10 is a collection of machine learning algorithms for data mining tasks. The classifiers in WEKA 3.6.10 were methods for predicting group membership based on numeric values. The experiment constructed one decision tree for each of the three image types.

Research Methods

This dissertation researcher enhanced categorization of the images by using image

registration and subtraction, and applying global and local thresholding methods. Global

thresholding used one threshold over the entire image to compute the image, while in

local thresholding the image was broken down into smaller sub-regions, and each

threshold was computed individually. A thresholding method was considered point

dependent if the threshold selection was based only on the gray-level of each pixel. For

global thresholding, the researcher used the method of Otsu (1979) for optimum

thresholding, in which thresholding of a global region was determined from a local

property in the neighborhood of each pixel. Some region-dependent methods were

developed to improve the histogram of an image, making it easier to select the global

threshold value.

For local thresholding methods, the flexibility of quadtree-based algorithms was

combined with the non-shaped-restricted features of region-growing schemes. The

purpose of local thresholding was to preserve features or edges in a gray-level image that

were lost using global thresholding techniques. Global methods preserved only global

features. However, it was often necessary to enhance details over small areas. Because

these areas were small, they may have had a negligible influence on the calculation of

global threshold values. Therefore, the local details were usually lost by global

thresholding. Local thresholding techniques obtained the local details. Local

thresholding maximized edge features by thresholding local regions separately.

The regional growing algorithm that was introduced poses no restrictions on region

shape or size, and the researcher tailored it to allow constraints to be applied and

approximated with ease. The work of Miyake et al. (2009) was similar to this study, but

differs technically in that the researchers did not use a splitting process prior to merging-

and-thresholding. The threshold could be derived from:

1. The average brightness of a region.

2. The average of the minimum and maximum brightness present in the region.

3. K-mean clustering where K was some defined variable.

The researcher conducted all the experiments comprehensively using the Matlab®
Image Processing Toolkit (IPT) to compare the splitting and merging algorithms on
mammographic, gray-scale, and CT and MRI mammogram images. The algorithm
utilized here was selected for its combination of commitment, use of heuristic
information, and exhaustive nature.

The following were the independent variables and experiments that were run on all
combinations of the data:

1. Image Registration using two different image types: CT image vs. MRI
   image. The image was taken from databases that were backed by ACR.
   Matlab provided full image registration, which was part of Matlab's IPT
   toolbox. The IPT supported many image file types (Linh et al., 2010). The
   results were analyzed using accuracy of segmentation (Verbeek et al., 1988).

2. Temporal subtraction vs. dual-energy subtraction. This experiment compared
   temporal subtraction using images from the ACR database, and applied
   temporal subtraction using the Matlab IPC toolbox and dual-energy
   subtraction images that were used from Lewin's (2003) research.

3. Eight different types of thresholding methods were run on all combinations of
   temporal and dual-energy images listed below.

   - Amplitude thresholding in Matlab. The Matlab function could be used to

find local minima, maxima, and match value in the vector of an image.

Matlab had many examples on the vector of an image (Matlab, 2012).

- Global thresholding. In this experiment both temporal subtraction and dual-energy subtraction images were used with DIPUM Toolbox Version 2 (Gonzalez, Woods, & Eddins, 2012). The DIPUM toolbox provided a function for the global image thresholding method. This experiment's analysis used the accuracy of segmentation and quality of classification.

- Local thresholding DIPUM Toolbox Version 2 (Gonzalez et al., 2012) provided a local image thresholding method. The image came from the experiment image registration using two different image types: CT image vs. MRI image. Analysis for accuracy of segmentation and quality of classification was completed.

- Connected-component labeling. Matlab had some examples for connected-component labeling located in the file exchange (Matlab, 2012). This experiment completed the analysis for the accuracy of segmentation and quality of classification.

- Threshold with histogram analysis. Matlab had a few functions for image histograms in the IPC that this dissertation researcher used in this experiment (Matlab, 2012).

- Thresholding with K-means clustering where K is some variable. Matlab had some examples for K-means clustering located in the file exchange. Analysis for the accuracy of segmentation and quality of classification was done on this experiment.

- Regional growing thresholding algorithm. This was obtained from the Gonzalez et al. (2012) function (regiongrow). An analysis for the accuracy of segmentation and quality of classification was used on regional growing completed image.

- Split-and-merge method thresholding. This was obtained from the Gonzalez et al. (2012) function (splitmerge). Analysis of the accuracy of segmentation and quality of classification was used.

4. WEKA 3.6.10 was used for obtaining precision, recall, and f-measure. The training was derived from the collection of 30 images containing two classifications—malignant and benign.

Resources

This experiment used the University of South Florida's mammography database (University of South Florida, 2011) and the American College of Radiology's breast imaging (BI-RADS) database. These databases were the standards for American College of Radiology's breast imaging (Vasantha & Bharathi, 2011).

Mammography image results from each of the experiments were collected and analyzed with the following data results.

1. Accuracy of segmentation via thresholding depended almost entirely on the selection of the value for the threshold. Good threshold values were estimated by using statistical information about the image like the average, median, and mode of the pixels' brightness in the image. Recall from the thresholding section was that first-order statistics were the accuracy of segmentation via thresholding; this depended almost entirely on the selection of the value for

the threshold (Chen & Chen, 2006). The average, median, and mode of the pixels' brightness were examples of these (Verbeek et al., 1988). By using different threshold methods one can compare accuracy and precision of thresholding against the pre-defined cancer images.

2. Quality of classification often is measured by precision, recall, and f-measure (which combines precision and recall) (Dembczynski, Waegeman, Cheng, & Hullermeier, 2011). These measures were applied to assess the relative effectiveness of these independent variables (i.e., subtraction method, segmentation method). All measurements were used to assess the effectiveness of the proposed methods (Kulkarni & Nicolls, 2009).

Throughout the study, many resources were used, including material from the following: Institute of Electrical and Electronics Engineers (IEEE), Association for Computing Machinery (ACM), Signal Processing Society, Society of Photo-Optical Instrumentation Engineers (SPIE), and Nova Southeastern University online library. The framework utilized Matlab and Simulink R2011b Environment applications. This was an open environment for developers and allows for C and C++ interface calls (Matlab, 2012). The Image Processing Toolbox extended Matlab and provided a complete set of references, standard algorithms, and graphical tools for image processing. The Image Processing Toolbox supported different types of image file formats for medical images, such as tomographic images (Matlab, 2012). The Image Processing Toolbox also provided comprehensive image display capabilities such as zoom, pan, and examination of regions of pixels. The GUI interface allowed interactively looking at Region of Interest (ROI), adjusting contrast, cropping, and measuring distance (Matlab, 2012).

Additional methods that were used within Matlab were from functions that could be found in the DIPUM Toolbox Version 2 (Gonzalez et al., 2012). The DIPUM Toolbox contains Matlab functions that Gonzalezet et al. used in their 2004 through 2012 textbooks.

Summary

The goal of this research was to assess the effectiveness of various thresholding methods in the context of a three-stage approach to help radiologists find cancerous tissue lesions in CT and MRI mammography images, using image registration and subtraction, thresholding, and classification for the detection of cancerous tissue with a CAD system design. There were two independent variables: subtraction method in stage one (temporal subtraction and dual-energy subtraction), and thresholding method in stage two. The remaining factors were fixed (including the classification method). All the experiments used the Matlab R2012a Environment.

The available software comes from the Matlab Image Toolbox and the DIPUM Toolbox. The built-in functions from the Matlab Image Toolbox were registration, image subtraction, region of interest, visualization, and algorithm development (Matlab, 2012). The Otsu (1979) function was available in Matlab and the DIPUM Toolbox. The researcher presents and compares the results image by image in a tabular format in Chapter 4.

Chapter 4

Results

Organization

This experimental design presents the methodology for using eight distinct thresholding methods to find the region of interest (ROI) containing malignancy cancer cells on three different image types: temporal subtraction, dual-energy subtraction, and Digital Database for Screening Mammography (DDSM). Ten unique images exist per type. In each of the three different image types, malignant cancer cells have been identified by radiologists using an image mask. From each image within the three image types, two classifications are created as follows: sub-regions from the image as malignant and benign. *Malignancy* refers to the part of the image that contains cancerous cells that have the ability to spread to other areas within the body (Dromain et al., 2012). *Benign* refers to the part of the image that has no cancerous cells found (Vasantha & Bharathi, 2011). This process yields 20 training instances in total for each image type.

The Matlab 2012a® Image Processing Toolkit (IPT) function Imtool has several built-in tools, including one that helps measure distance in an image for determining the precision value. Imtool also allows for cropping an area of interest and saving the cropped image. Using the Matlab 2012a functions co-occurrence matrix (GLCM) and histogram based features, 25 different feature types are captured. GLCM calculates 21 target features types which are obtained from the Matlab function texture features. The remaining four feature types—mean, variance, skewness, and kurtosis—are obtained

from the state of the histogram based features.

Using WEKA® 3.6.10, a training set was derived from the collection of 30 images containing the two classifications—malignant and benign. The test set was derived from thresholding each of the three different image types (temporal subtraction, dual-energy subtraction, and DDSM) running the feature types on each of the two classifications. The final step was to then run the test set derived from applying a thresholding method to a single image against the training set decision tree, which provides the precision, recall, and f-measure values. This methodology forms an assessment for the combination of image type and thresholding method, which yields the calculated precision, recall, and f-measure values. The rest of this section elaborates the steps in this experimental methodology.

Image Data

The images data were from three sources. The first set of images was taken from the Digital Database for Screening Mammography (DDSM) from the University of South Florida's mammography database. The second set of images, temporal image subtraction, came from two sources. First, from DDSM case 83 and case 84 (referenced as Patient A, Figure 18) and DDSM case 171 and case 216 (referenced as Patient B, Figure 19) were created using Matlab IPT image registration. The remainder of the temporal image subtraction images are from Rafferty (2007). The third set, dual-energy images, came from Lewin's (2003) research. Patient 2 was referenced as Patient C, Figure 21, and Patient 6 was referenced as Patient D, Figure 21.

The dissertation researcher investigated three types of image data: temporal subtraction, dual-energy subtraction, and DDSM. DDSM contains descriptions of breast

lesions in terms of the American College of Radiology (ACR) and breast imaging lexicon

called the Breast Imaging Reporting and Data System (BI-RADS) (Vasantha & Bharathi,

2011). Temporal subtraction has been used successfully for the detection of lung nodules

by Miyake et al. (2009). Carton et al. (2007) used dual-energy subtraction to find cancer.

One issue that the researcher had to resolve was data had to be transformed to standard

grayscale image. This was seen in the predominant negative aspect working with

temporal and dual-energy images, as data were presented with RGB grayscale values (0–

255). The RGB color scheme allows 256 values that are possible for shades of color.

These values have to be converted or transformed to equal standard grayscale image

values (0–255) in order to prepare for subsequent stages of the experiments.
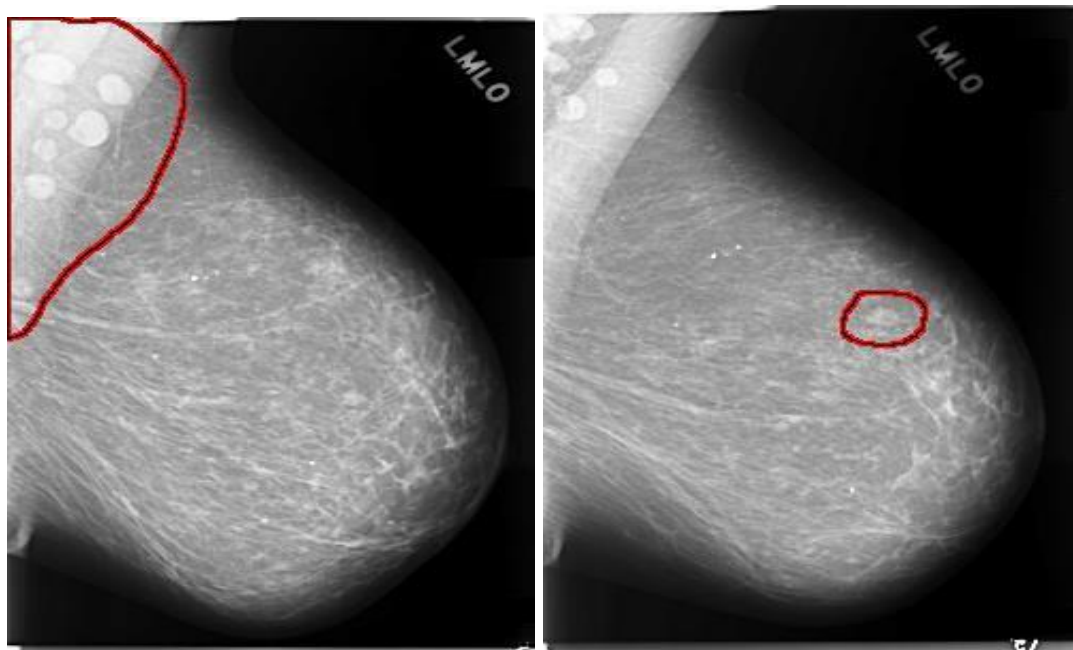
Figure 18. Images from University of South Florida Digital Mammography (DDSM),
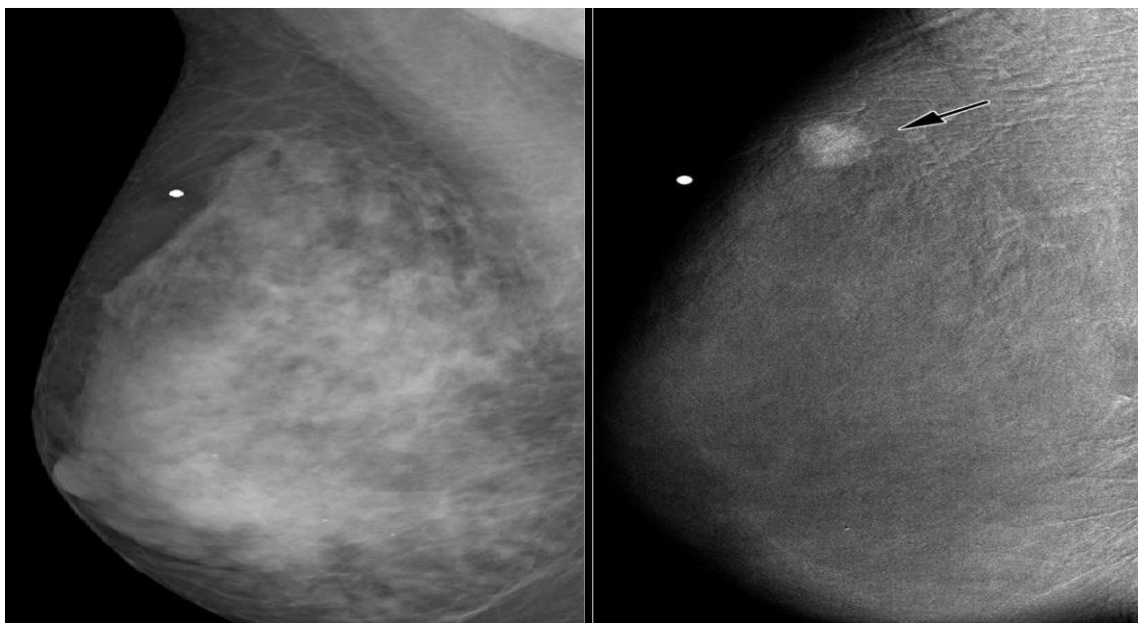Patient A.
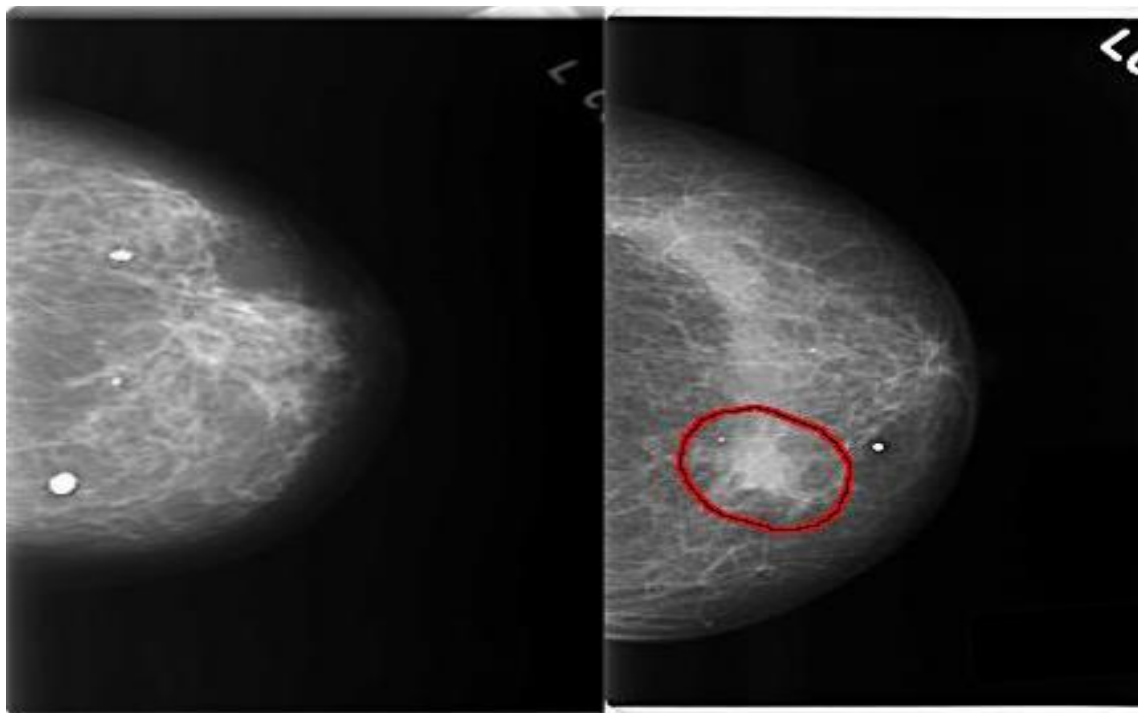
Figure 19. Images from University of South Florida Digital Mammography (DDSM), Patient B.

(a)   (b)

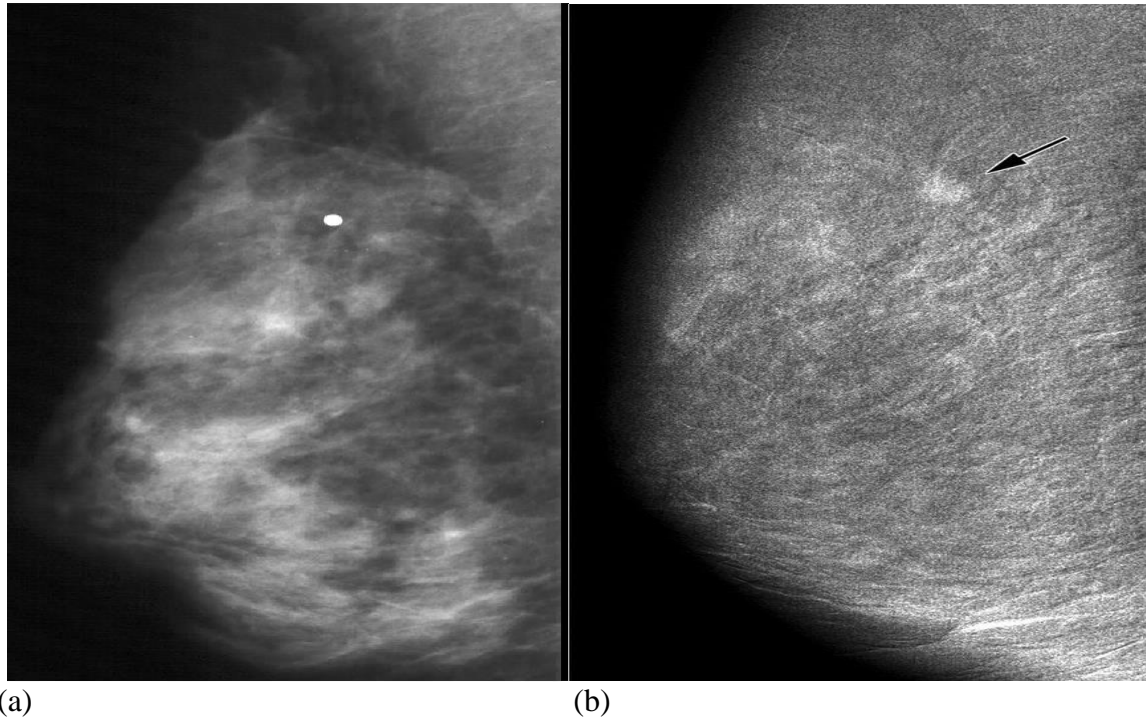*Figure 20*. Lewin (2003) images Patient C; (a) Benign dual-energy.

(a)   (b)

Figure 21. Lewin (2003) images Patient D, (a) Benign dual-energy.

Deriving Training Sets

For deriving the training, sets the resarcher looked at classifying feature vectors of an image based on the Region of Interest (ROI) in the image (Vasantha & Bharathi, 2011). Both original and thresholded images have an ROI of the malignant region that had been previously identified by a radiologist. The extraction of the ROI was necessary to analyze the feature vector, and to find the prominent features that were representative of the classes of the images. *Malignancy* refers to cancerous cells that had the ability to spread to other areas within the body (Dromain et al., 2012). All malignancy had been identified by radiologists in each image type using an image mask to identify the ROI malignancy. *Benign* refers to a sub-image that has no cancer found (Vasantha & Bharathi, 2011). Note that the benign and malignant regions of an image partition the image, and combine to form the Image as a whole. The benign area is all parts of the

original image other than the portion containing the malignancy that was identified by the

radiologists in the foreground of the whole image. When applying thresholding

techniques to the image, either the original radiologist's mask overlay or the original

image mask to identify the area of malignancy. To summarize, the structure of the

training sets are derived from the whole image, and the sub-images identified by the

radiologists' designation as malignancy and the remainder, benign, found in the area

other than the malignancy.

The Matlab 2012a image processing toolbox function Imtool has a built-in tool that

helps measure distance in the image for precision. Imtool also allows for cropping an

area of interest. To crop the benign area, all areas other than the malignancy that was

identified by the radiologists in the foreground were selected. The Matlab 2012a

supported feature set GLCM and histogram were run on each of the image

classifications—malignant and benign. This provided 30 images along with their

corresponding thresholding methods and feature sets, which provided an adequate

representation in the training set.

Constructing the Decision Trees

Each image has two parts for classification: malignancy refers to the part of the image

that contains cancerous cells and benign refers to the part of the image that has no

cancerous cells. For image classification, the dissertation researcher used a J48 decision

tree on 21 features that were identified and obtained from the Matlab GLCM function and

four feature types that were obtained from the state of the histogram-based features.

WEKA 3.6.10, along with cross-validation, was used to create a decision tree training set

using the 25 features that were relevant at each node of the two. The feature vector for a

region was obtained from the original thresholded whole image.

There are three different image types: temporal subtraction, dual-energy subtraction, and Digital Database for Screening Mammography (DDSM). The classification used for the construction of decision trees supports malignant and benign. The feature vectors that are based on malignant regions were used on malignant instances and the feature vectors based on benign regions were used for benign instances. The image was based on the image as a whole and the specific region being examined for the feature vectors.

WEKA 3.6.10 was used with the J48 classifier and cross-validation to create the final decision trees. Vasantha and Bharathi (2011) used the Machine Learning package WEKA to train the dataset, and also used a decision tree method. WEKA 3.6.10 is a collection of machine learning algorithms for data mining tasks. The classifiers in WEKA 3.6.10 were methods for predicting group membership based on numeric values. The experiment used one decision tree for each of the three image types (temporal subtraction, dual-energy subtraction, and DDSM). For creating the decision tree, the feature vectors and classifiers are malignant and benign regions from the 10 cases.

Deriving Test Sets

In the test set, the researcher derived an ROI obtained in the thresholded image, generated a feature vector for this ROI, and applied this vector to the relevant decision tree. The test sets are derived from applying four different thresholding techniques. The data used were 30 images containing malignancies that were comprised of 10 DDSM mammographic images, 10 dual-energy subtraction images, and 10 temporal subtraction mammogram images. The approaches for image thresholding fall into eight broad categories: amplitude thresholding, global thresholding, local thresholding, connected-

component labeling, thresholding with histogram, k-means clustering, regional growing thresholding, and split-and-merge thresholding. The histogram thresholding method used was part of the Matlab-supported feature set and constructed four feature types: mean, variance, skewness, and kurtosis. These were used along with the 21 GLCM target feature types to construct the decision trees for each image type and also to construct test sets for thresholding.

The same Matlab 2012a supported feature set GLCM and histogram values were applied on the same malignancy and benign sub-regions that were extracted to create the training set. For each of the three image types, a thresholding method was applied. Hence, each test set is characterized by a sub-image type, threshold-method pair. Each test set comprised the set of 10 images of a given image type subject to the given threshold method. The ROIs determined by thresholding were used to derive the feature vector, and to find the prominent features that were representative of the classes of the thresholded image. The test set was used to verify an ROI whose vector would be classified as malignancy under the decision tree for the image type. This process provided an individual test set that could be run against the full training set using WEKA 3.6.10.

Classification

The ROI was determined by the thresholding method being tested. The researcher derived a feature vector for the ROI by thresholding an original image that classified as malignant, and a feature vector for the complement of the ROI in an image, which was any other area in the foreground that was classified as benign. Identifying the ROI after thresholding is a subjective measurement, as the original image has been transformed by

the selected thresholding method.

Both positive and negative test cases were used in the experiments: the positive malignancy cases were identified by thresholding the original image, and the negative benign cases were complements of the malignancy region. The feature vector for a region was obtained from the original un-thresholded whole image. The experiment used one decision tree for each of the two image types and then classified both the malignant and benign regions from the 10 test cases.

The results of the experiments were used in the calculation of precision, recall, and f-measure. *Precision*, also called positive predictive value, is the fraction of retrieved instances that are relevant in the image (Russell & Norvig, 2003). *Recall* is the fraction of relevant instances that are retrieved in the image (Russell & Norvig, 2003). Finally, *f-measure* is the measurement that tests for accuracy in both precision and recall. For an experiment, precision and recall depends on true and false positives. The maximum precision would have no false positives and the maximum recall would have no false negatives (Russell & Norvig, 2003). An experiment is parameterized by image type and thresholding method.

The series of steps in this experiment provided for a methodical evaluation of each image type against each thresholding method, which was then used to create a test set that helped determine the adequacy of the individual methods. Ultimately, this methodology formed assessments for each combination of image type and thresholding method. The researcher examined the results to determine which combination yields the greatest calculated precision, recall, and f-measure values and provide a benchmark for comparing the results across each cell of the image type thresholding method matrix.

Findings

In the first stage, image registration was used for temporal subtraction and dual-energy subtraction. The fundamental characteristic of any image registration is the type of spatial transformation or mapping used to properly overlay two images (Matlab, 2012). Since a relatively high image contrast (mammograms) was the source, the Matlab registration process was the method used. The Matlab registration process allowed different tomographic images (MR, CT, PET, etc.) when the size of the pixel was known.

Image registration used feature identification to find some information that was invariant to the registration process. The information found to be the same in all mammograms involved the external boundaries. Regardless of the type of image (MR, CT, PET, etc.), as long as the subject was the same patient, and no significant external shape modifications existed, the boundaries should have had the same shape. This dissertation researcher determined a set of features completely automatically. Since this method was designed to work with any kind of subject data, this step's approach used gray-level information. Figure 22 shows the registration interface on the left-hand side with the two images that were registered for Patient A. In the center, the result of the image registration image is presented. Figure 23 presents registration using temporal subtraction for Patient B.
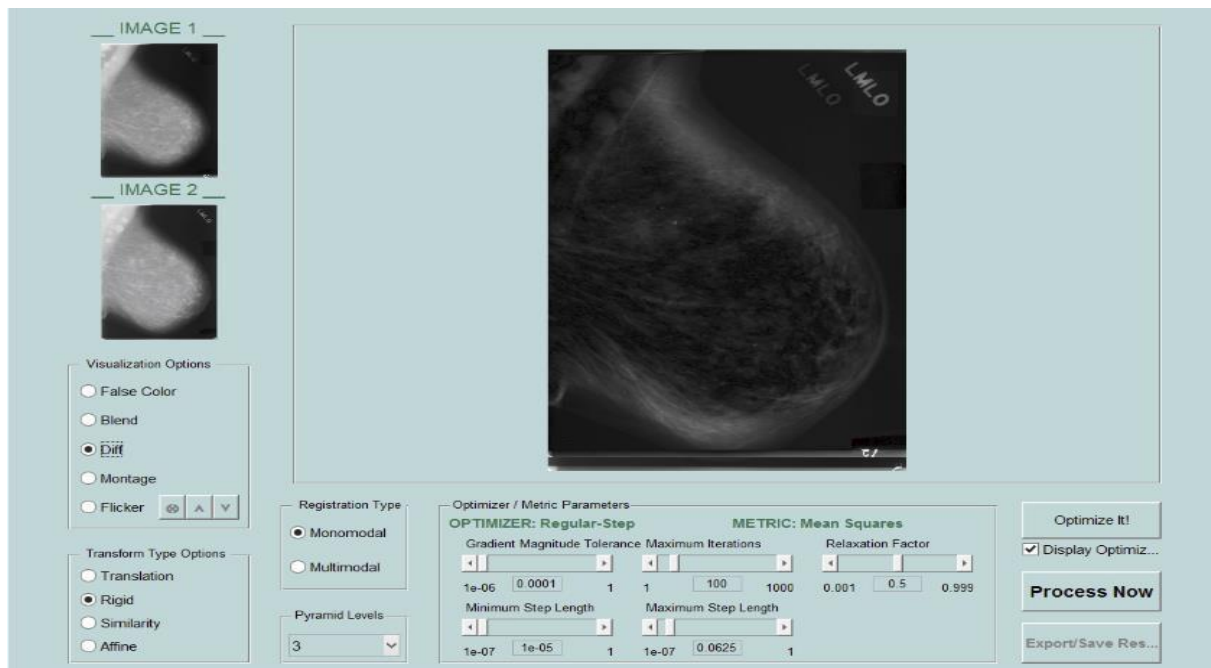
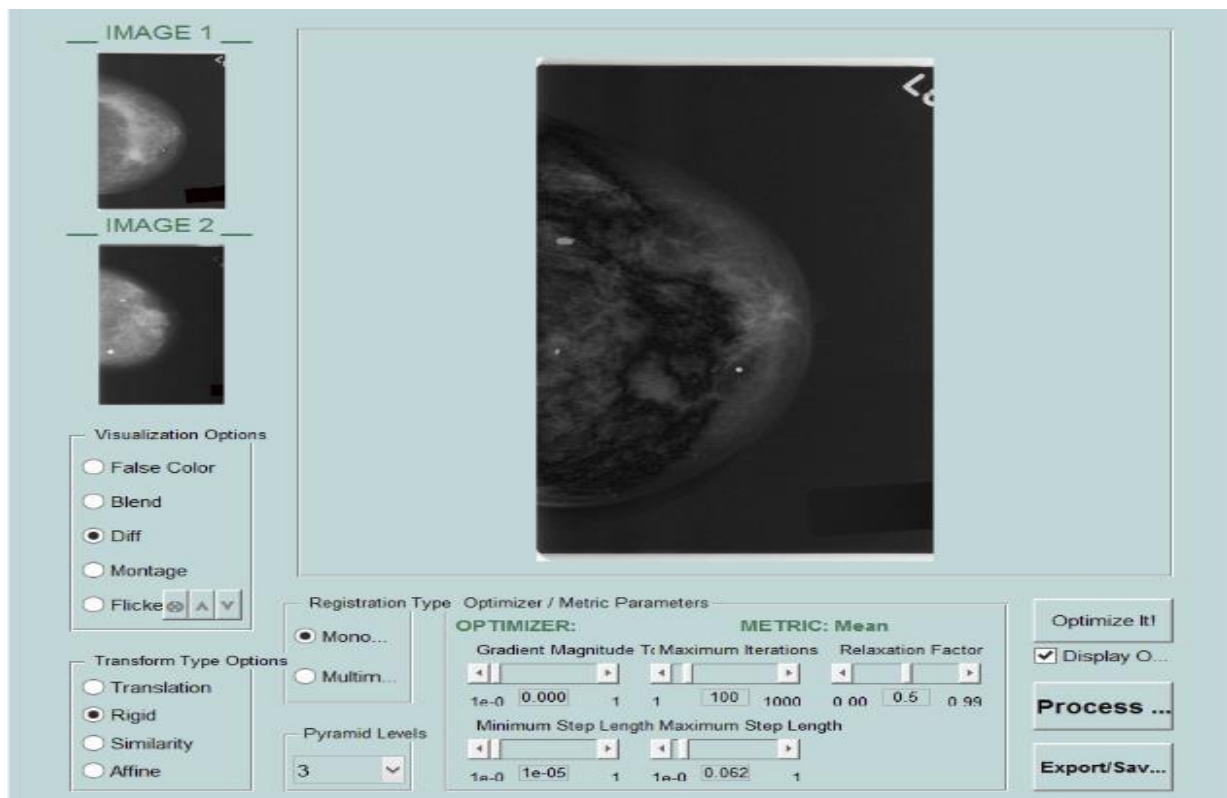*Figure 22*. Image registration using temporal subtraction Patient A.

Figure 23. Image registration using temporal subtraction Patient B.

The researcher accomplished the second stage by applying thresholding techniques to find the Region of Interest (ROI). When the ROI is identical to or similar enough to the one identified by the radiologist before thresholding, it is described as *holding the ROI* (Miyake et al., 2009). This term is used to identify a region that was noted to have a malignancy, which refers to the area of the image that contains cancerous cells. If the ROI is very similar but not identical to the original ROI, this would be an area that holds the ROI because thresholding will change the look but not the overall size and degrees of the area. Matlab's 2013 distance tool in the Imtool set allows for measuring distance in images by using the tool function in the toolbar or by selecting Measure Distance. This allows the distance between two endpoints to be calculated and expressed in data units determined by the XData and YData properties, which by default is in pixels. The Imtool allows for exporting endpoint distance data, thus allowing end point locations and distance information to be saved (Matlab, 2013). The researcher used this tool to measure the distance in each image both before and after thresholding the area defined by the radiologist. The Imtool is used to show the size of the identified malignancy ROI. The distance from the original reference point prior to thresholding is compared to the distance calculated after thresholding to determine if the new region holds the ROI, helping to validate the legitimacy of the other calculated values.

Thresholding is the process of finding individual pixels in an image that are greater than a specified threshold value. The researcher used eight different types of thresholding methods in this section. In amplitude thresholding, ground truth was noticeably different, and dual-energy seemed to hold the ROI identified by the radiologist, whereas temporal subtraction did not. ROI was defined by a radiologist on

all the different image types prior to any experiments and noted as malignant. Global

thresholding did not detect ROI in temporal subtraction or dual-energy. Local

thresholding was different in that dual-energy seemed to find the ROI better than

temporal subtraction when the image was broken down into smaller sub-regions and the

edge features were maximized in blocks. Connected-component labeling, like global

thresholding, did not hold the ROI. The only difference was in dual-energy; the image

looked grainy and the resultant image was black instead of white.

The researcher used histogram analysis thresholding to find four feature types that

were obtained from the state of the histogram-based features: mean, variance, skewness,

and kurtosis. K-means clustering seemed to do very well in both temporal subtraction

and dual-energy subtraction for finding ROI. The regional growing thresholding

algorithm seemed to be the best in the 4-neighbors function as it returned the ROI in both

temporal subtraction and dual-energy subtraction. The split-and-merge thresholding

algorithm did not do well on the split, as there was no sign of the ROI.

In the last stage, the researcher used a J48 classifier with cross-validation based on the

decision tree, which provided the precision, recall, and f-measure values. Classifiers

detect features based on shape and density in order to find malignant growth (Miyake et

al., 2009). This technique can provide a useful CAD application for feature detection of

malignant growth to help radiologists and physicians detect and diagnose possible

cancerous growths. The researcher identified and obtained 21 features from the Matlab

GLCM function, and four features were obtained from the histogram data. The WEKA

results collected included precision, recall, and f-measure.

Amplitude Thresholding in Matlab

The next several sections on thresholding provide anecdotal evidence about the various thresholding methods based on a limited number of images. Using Matlab on adaptive thresholding algorithm separated the foreground from the background with non-uniform illumination (Xiong, 2005). Single and multiple amplitude thresholding techniques generated thresholds based upon the statistical features of the region. Figure 24(b) displays the resultant temporal subtraction images, and Figure 25(b) shows the dual-energy subtraction image results. The histogram of a digital image provided more information with minima and maxima of the image, and is discussed further in a later section of this chapter. One thing that was noticeable between temporal subtraction and dual-energy images in amplitude thresholding was that the ground truth temporal subtraction image was white, and the dual-energy subtraction image was black. Appendices A, B, and C are the results of each anecdotal thresholding method from each image type that consists of 30 images containing malignancies that were comprised of 10 DDSM mammographic images, 10 dual-energy subtraction images, and 10 temporal subtraction mammogram images. The researcher examined the results shown in the Appendices to determine which combination yielded the greatest calculated precision, recall, and f-measure values.
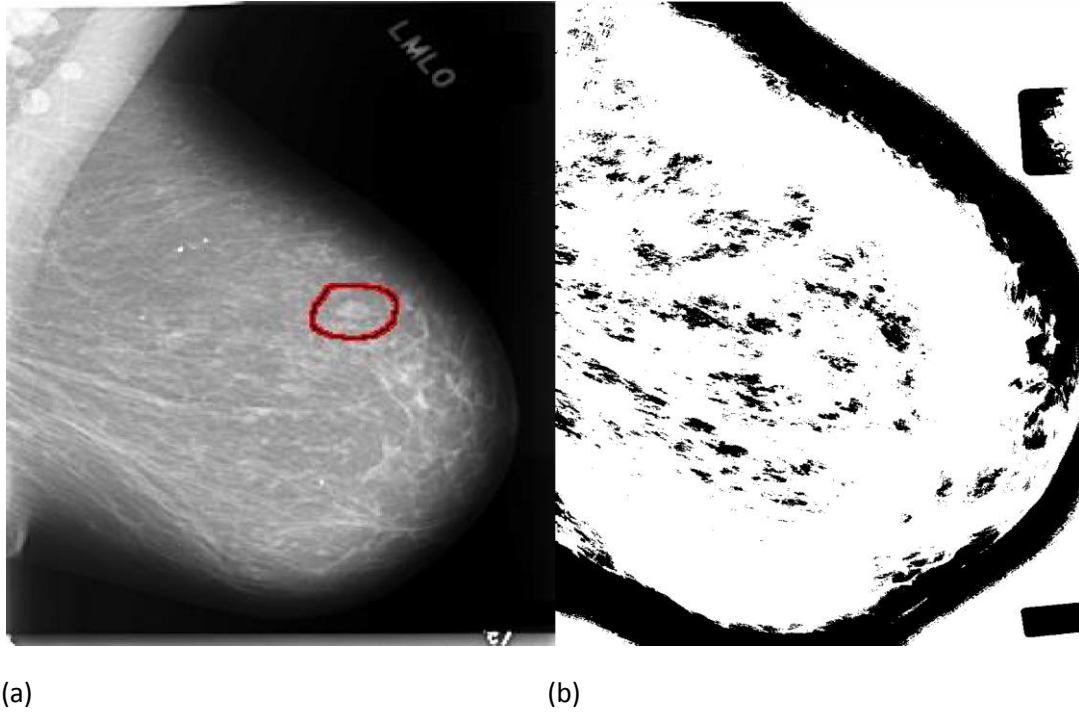
(a)   (b)

Figure 24. Amplitude thresholding using temporal subtraction image Patient A.

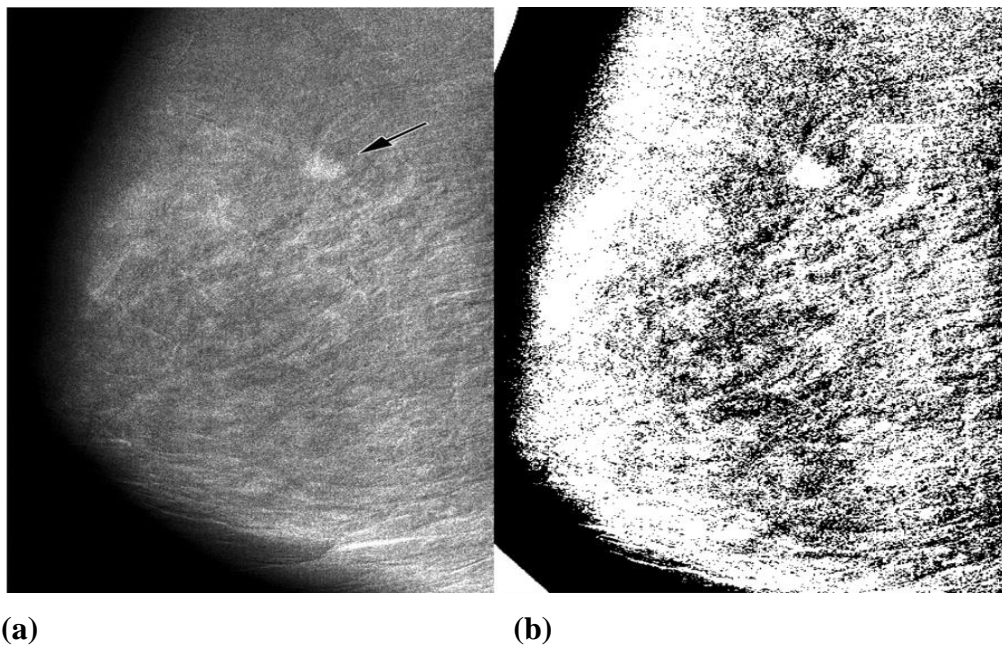**(a)**   **(b)**

Figure 25. Amplitude thresholding using dual-energy subtraction image Patient D.

Global Thresholding

In Global thresholding, the thresholded images in Figure 26(b) (using temporal subtraction), and Figure 27(b) (using dual-energy subtraction), contained all of the edges within the original images shown in Figures 18 and 20. In comparison to the global techniques, no portion of the image was missing in the thresholded image. The Otsu (1979) and the global contour methods did not detect one of the cancerous tissues shown in Figure 18 or Figure 20, respectively. Global thresholding did show the full contour of the image and followed the Otsu method. The global thresholding, connected-component labeling, and split-and-merge methods were not noted in the Appendices because all results were the same because all of the images had the same brightness level in the foreground of each image.
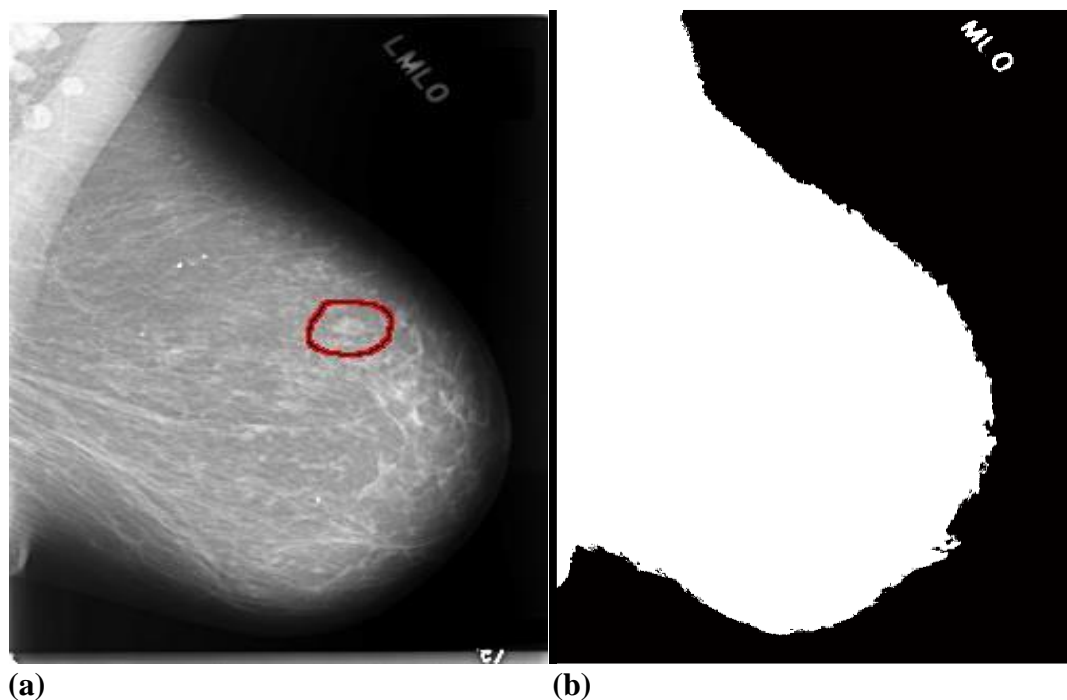

(a)   **(b)**

Figure 26. Global thresholding using temporal subtraction image Patient A.

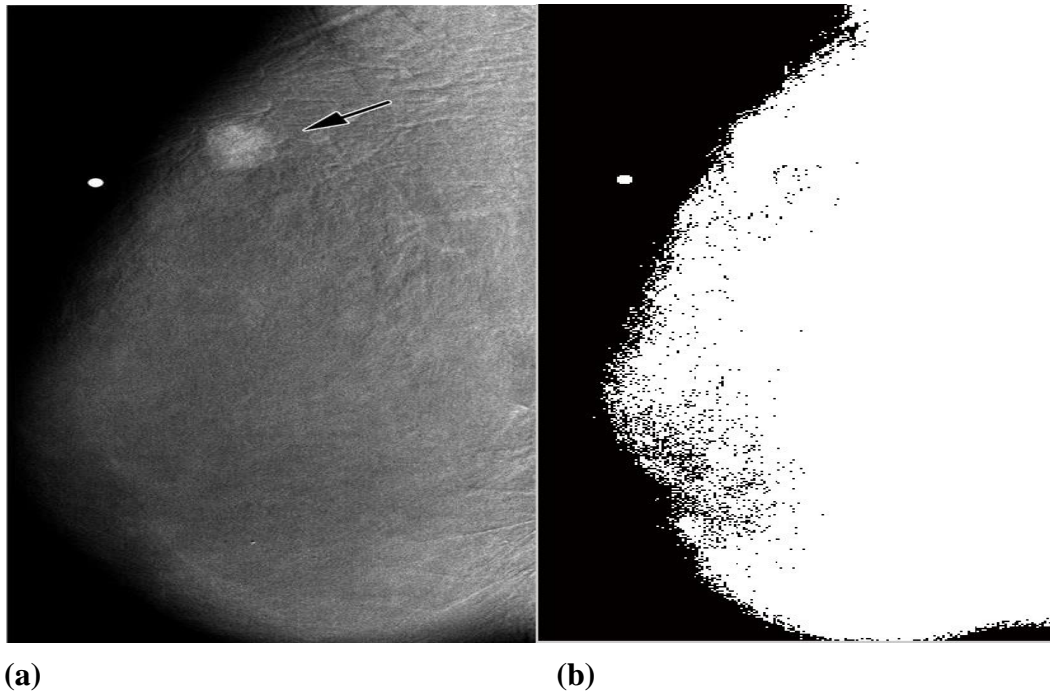**(a)**   **(b)**

Figure 27. Global thresholding using dual-energy subtraction image Patient C, local thresholding.

The local thresholding techniques presented in this dissertation were demonstrated using the same images as previously used for the global thresholding techniques. The results are shown in Figure 28(b) (temporal subtraction image) and Figure 29(b) (dual-energy subtraction image). Local thresholding processed an image, called the *graythresh function*, on local blocks of the image (Chow & Kaneko, 1972). This facilitated thresholding of an image with uneven background illumination, for which global thresholding was inadequate. This local thresholding process used the Matlab Image Processing Toolbox function.
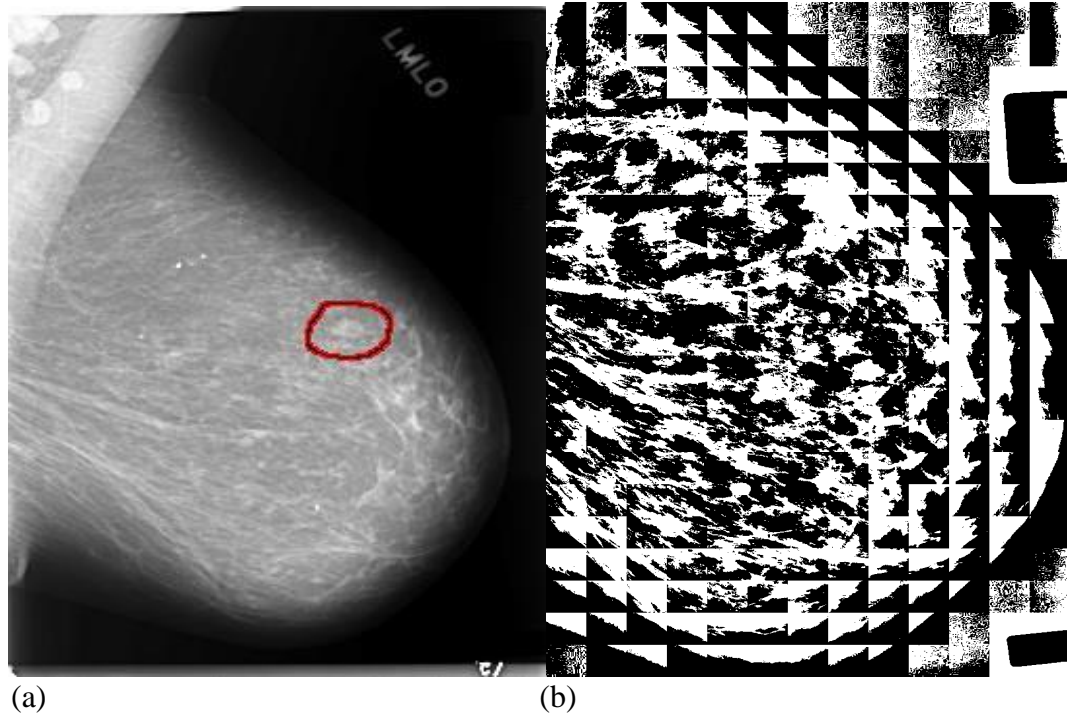
Figure 28. Local thresholding using temporal subtraction image; (a) Original image, and (b) Local thresholding, Patient A.
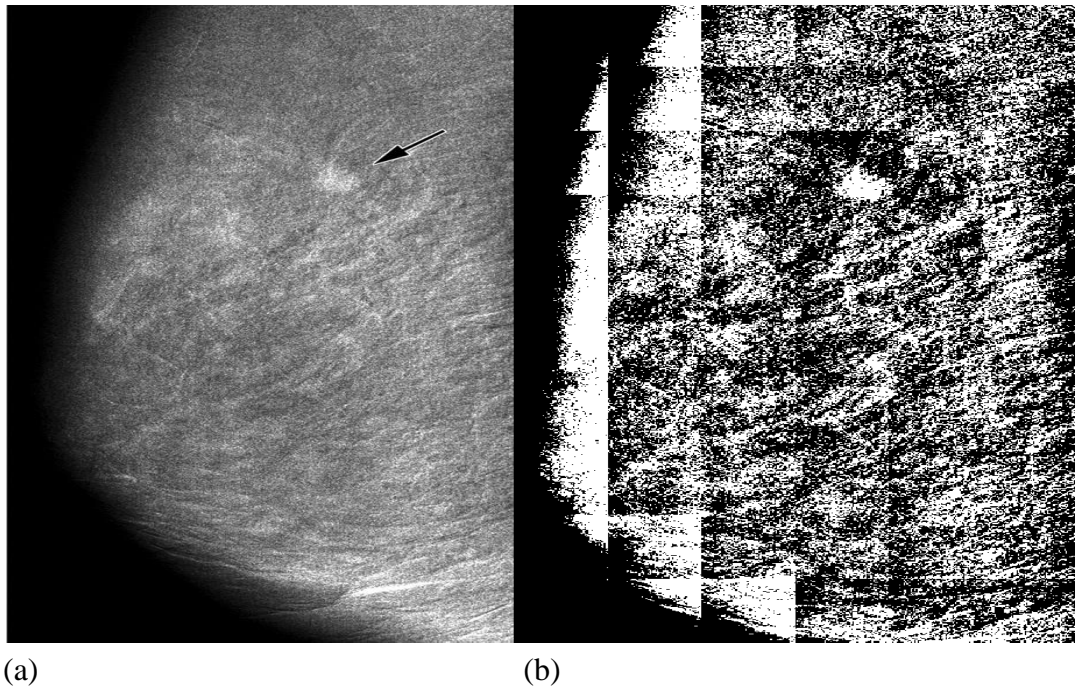

Figure 29. Local Thresholding using dual-energy subtraction image; (a) Original image, and (b) Local thresholding, Patient D.

As discussed in Chapter 2, in local thresholding, the original image was broken down into smaller sub-regions, and each was given a threshold individually. In many cases, this caused discontinuities at the borders between sub-regions. This effect can be seen in Figure 28(b) (temporal subtraction image), and Figure 29(b) (dual-energy subtraction image). See the results in Appendices A, B, and C for which the WEKA decision tree yielded the greatest calculated precision, recall, and f-measure values.

Connected-Component Labeling

A connected-component labeling algorithm may be used in conjunction with thresholding to satisfy the connectedness condition, as seen in Figure 30(b) (temporal subtraction image) and Figure 31(b) (dual-energy subtraction image). The results were similar to Figure 26(b) and Figure 27(b) representing global thresholding.
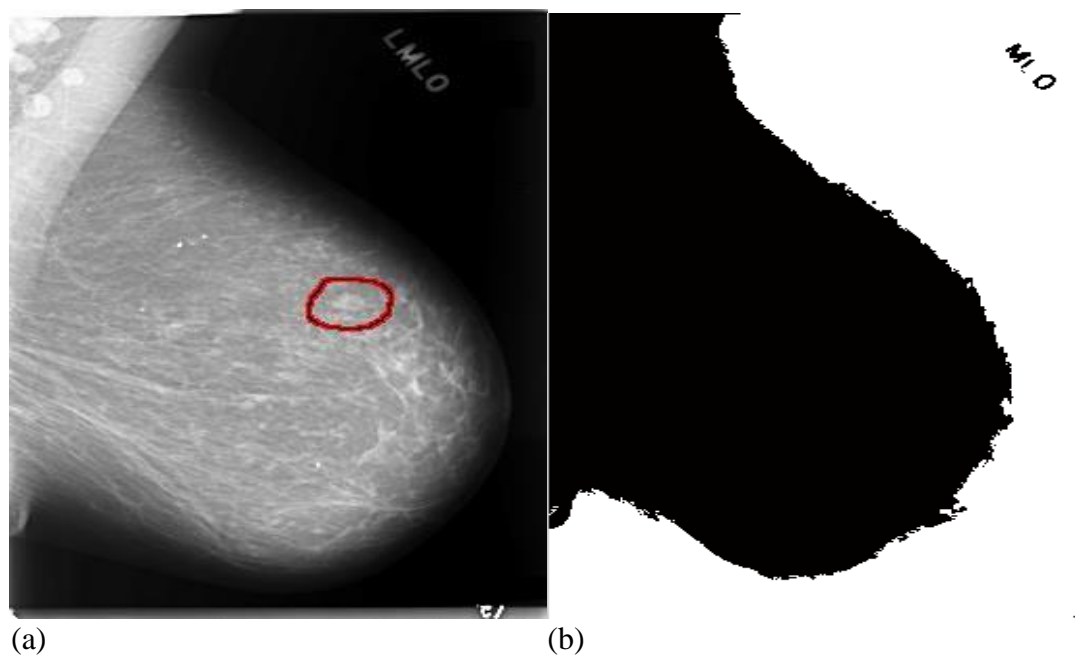

(a)   (b)

Figure 30. Connected-component labeling using temporal subtraction image, Patient A.

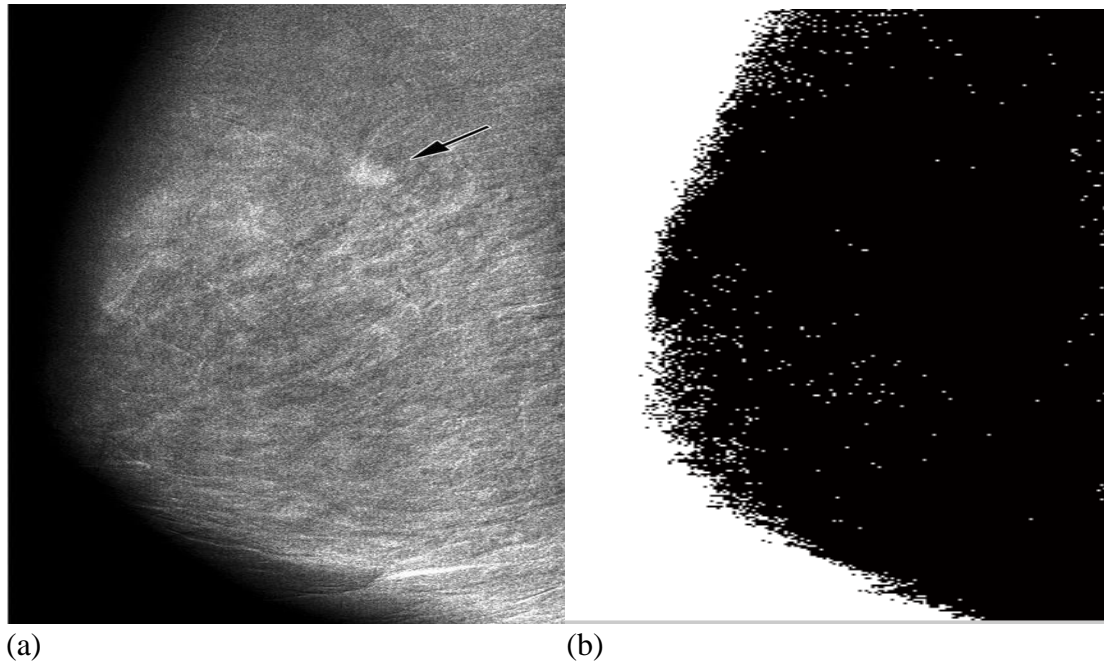(a)   (b)

Figure 31. Connected-component labeling using dual-energy subtraction image, Patient D.

Threshold with Histogram Analysis

The results of thresholding with histogram analysis can be seen in Figure 32, with the temporal subtraction image, and Figure 33 with the dual-energy image. Matlab had a few functions for image histograms in the IPC toolbox that the dissertation researcher used in this experiment (Matlab, 2012). The "imhist" function from the IPC toolbox displays a histogram for an image, and the best way to find the threshold value was to look at the histogram of the image. The researcher used the histograms later in this chapter in the classification section to find features mean, variance, skewness, and kurtosis. The energy and entropy values generated using this method were identical to their corresponding GLCM values.
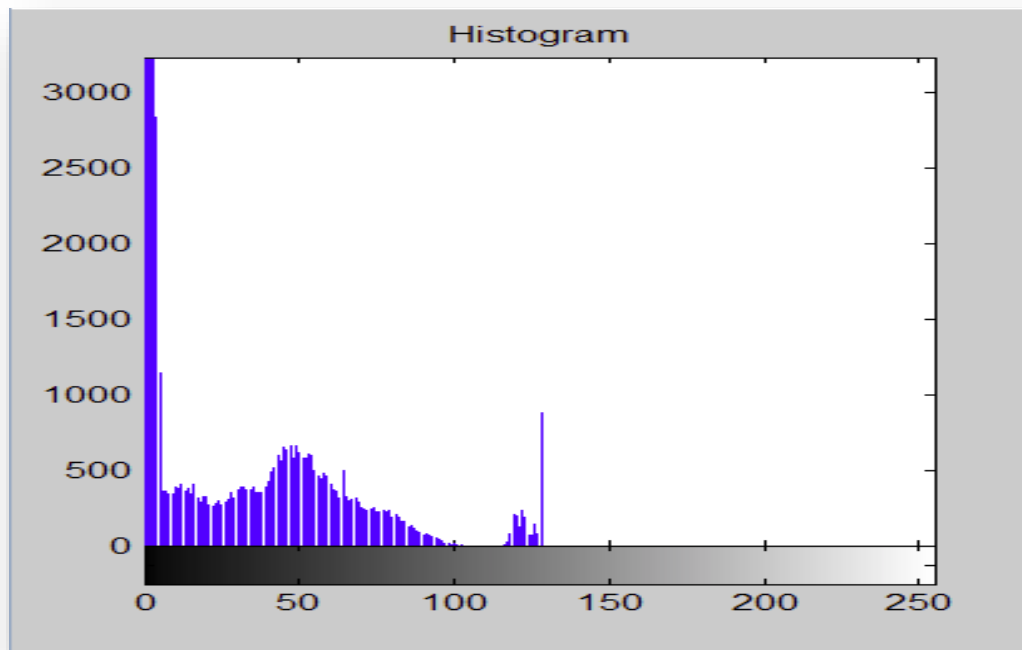
Figure 32. Threshold with histogram analysis using temporal subtraction image, Patient B.
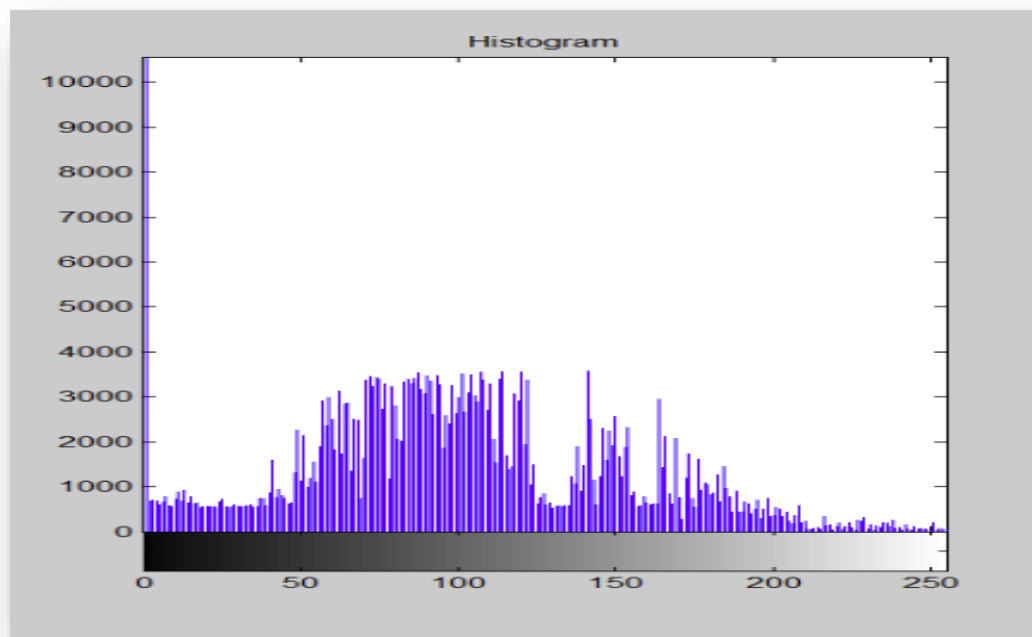

Figure 33. Threshold with histogram analysis using dual-energy subtraction image, Patient D.

Thresholding with K-Means Clustering

In the Matlab file exchange, the researcher found many examples of image processing with K-means clustering. The one used in this dissertation outputted the binary image and threshold level of the image using a 3-class, fuzzy k-means clustering, as shown in Figure 34 and Figure 35. Xiong (2005), the author of the function, explained it often worked better than Otsu's (1979) method, which output larger or smaller threshold levels. Xiong (2005) noted that K-mean clustering found the ROI at the higher levels of thresholding than was expected. In Figure 34, Using Temporal Subtraction Image, and Figure 35, Using Dual-energy Subtraction Image, it can be seen that the last images showed the ROI. See Appendices A, B, and C for the greatest calculated precision, recall, and f-measure values.
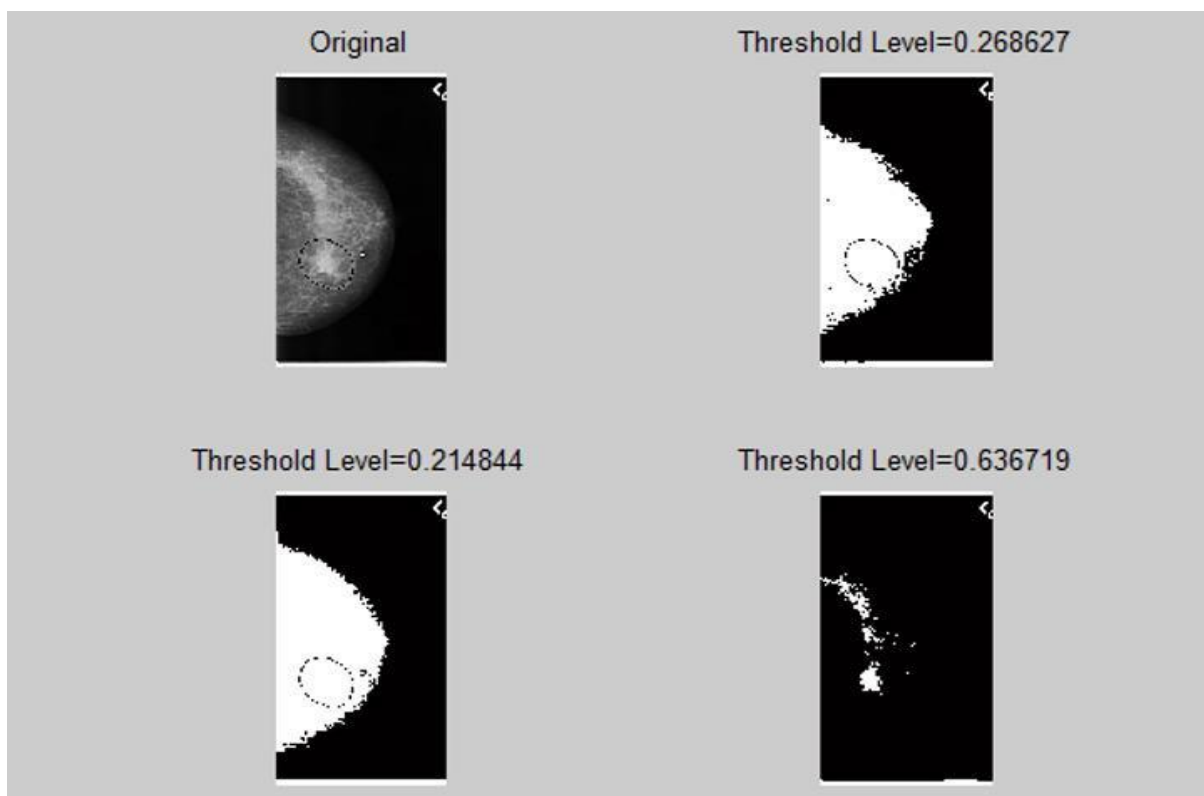

Figure 34. Thresholding with k-means clustering using temporal subtraction image, Patient B.
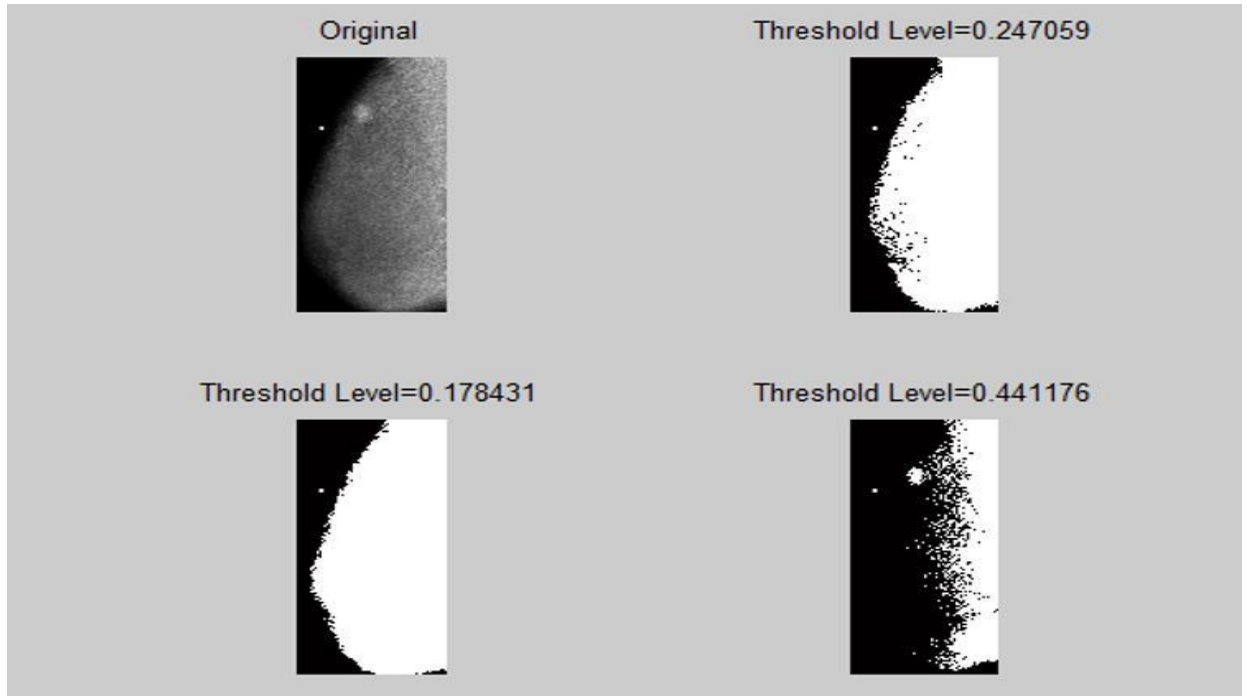
Figure 35. Thresholding with k-means clustering using dual-energy subtraction image, Patient C.

Regional Growing Thresholding Algorithm

The results of regional growing can be seen in Figure 36, using the temporal subtraction, and Figure 37, using the dual-energy subtraction images. The 4-neighbors function returned the coordinates of an element in a matrix for dual-energy, shown in Figure 37(b). The values were returned in the list of neighbors. If the neighbor did not exist, that was, if (x,y) corresponded to an edge or corner of the array, the list of neighbors contained only those coordinates corresponding to real neighbors.

The 8-neighbors function returned the coordinates of the element in a matrix just like the 4-neighbors function (Kroon, 2008), and seems to work well in temporal subtraction as shown in Figure 36(c). If the neighbor did not exist, that was, if the center corresponded to an edge or corner of the array, the list of neighbors contained only those coordinates corresponding to real neighbors. See Appendices A, B, and C for precision,

recall, and f-measure values.

(a)   (b)   (c)
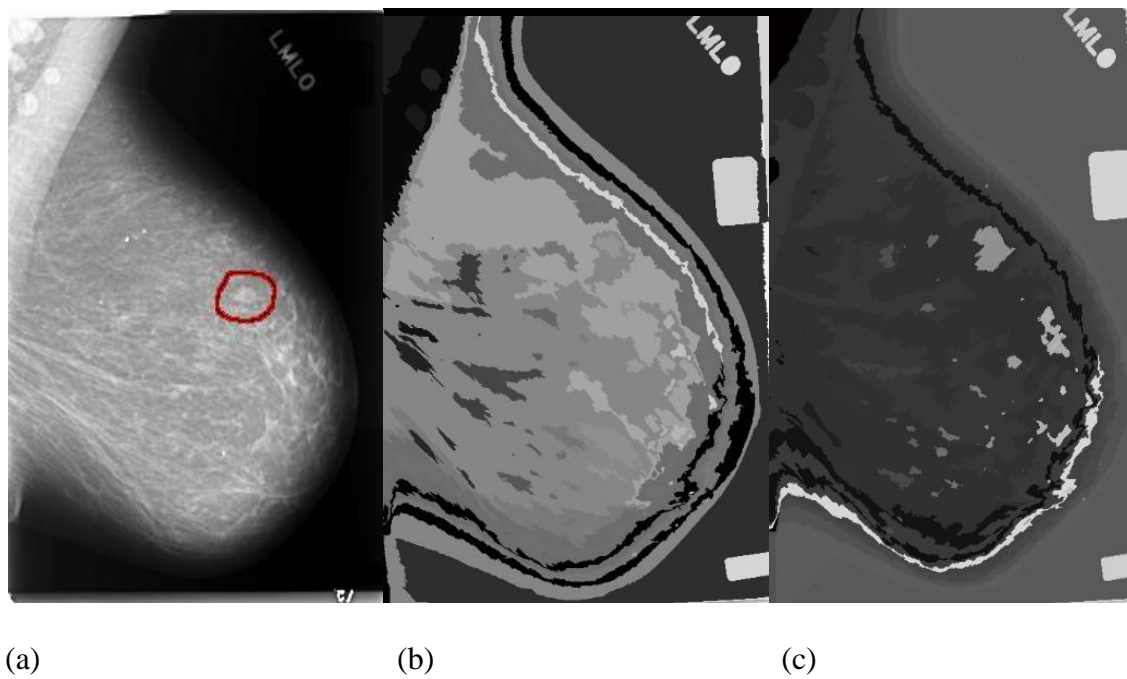
Figure 36. Regional growing using temporal subtraction image, Patient A

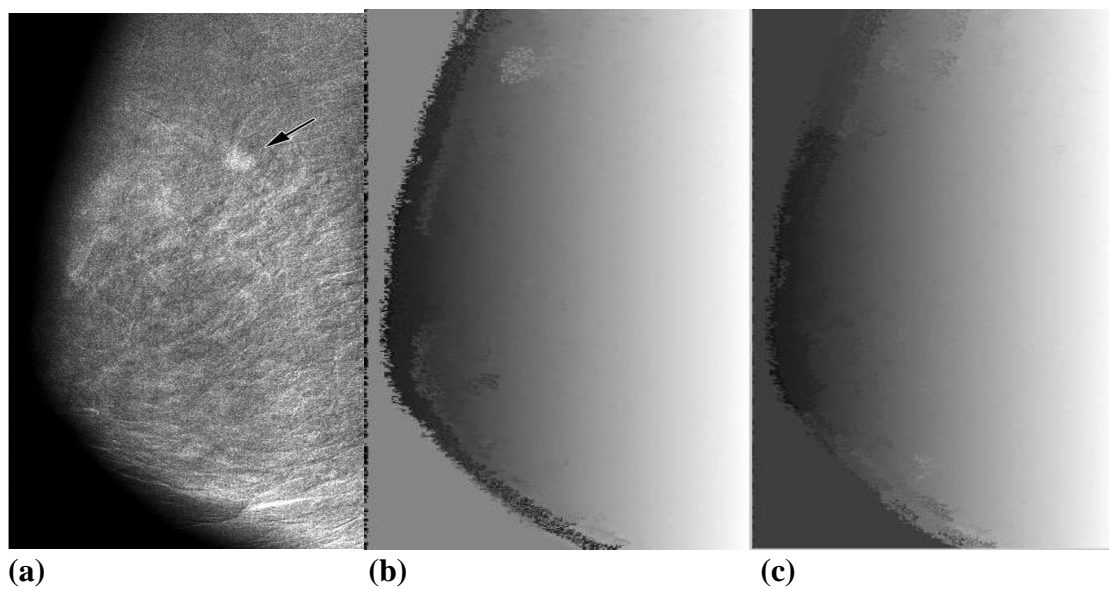**(a)**   **(b)**   **(c)**

Figure 37. Regional growing using dual-energy subtraction image, Patient D.

Split-and-Merge Method Thresholding

Through this method, the researcher performed basic region splitting on an image and used the Matlab function region split. The researcher used the value of the similarity criterion to decide between splitting of segments or not; the set value of the first label to be used (typically 1) was primarily needed in the recursion, and was set to an empty matrix initially. The split image was enhanced when viewed in color; when viewed in grayscale the image contained unrecognizable fragments. Figure 38(b) shows a split threshold using a temporal subtraction image and Figure 39(b) shows a dual-energy subtraction image.
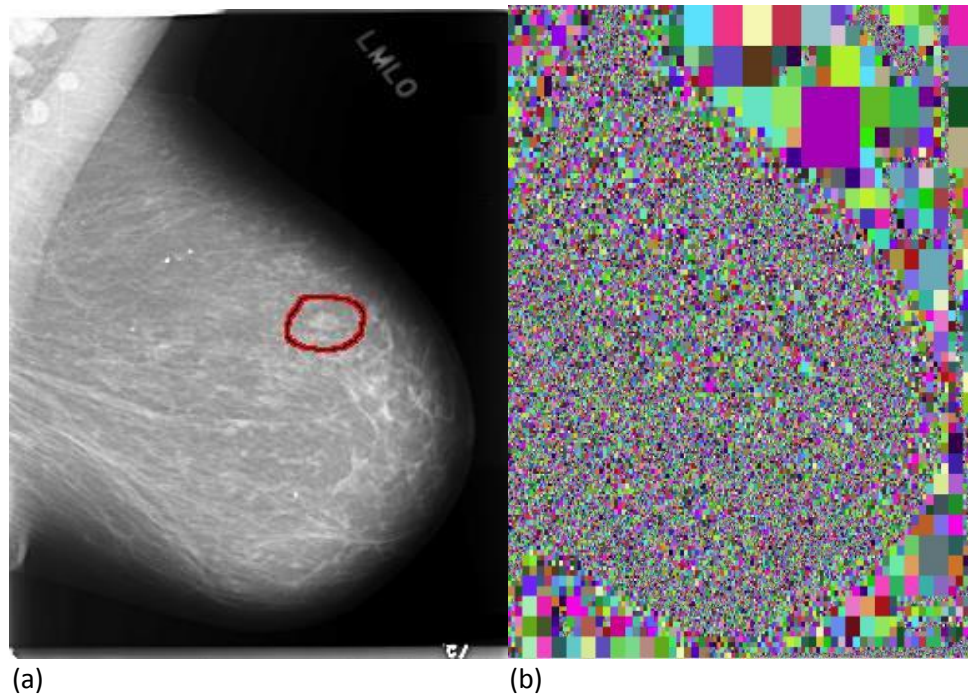

(a)   (b)

Figure 38. Split thresholding using temporal subtraction image, Patient A.

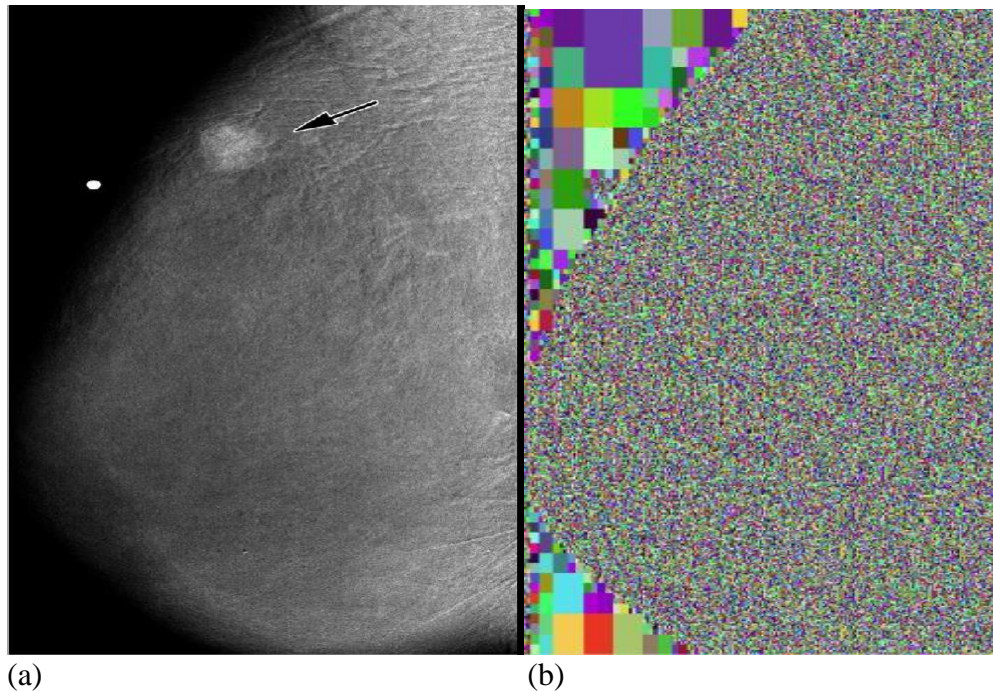(a)   (b)

Figure 39. Split thresholding using dual-energy subtraction image, Patient C.

Image Classification

   Classification is an important measurement in the areas of pattern recognition,

artificial intelligence, and vision analysis (Vasantha & Bharathi, 2011). These

measurements were applied to the independent variables of the subtraction method and

the segmentation method. Vasantha and Bharathi (2011) used 28 feature types. Twenty

one of the feature types used in this dissertation are the same as those used in the work of

Vasantha and Bharathi and the remaining four are from the histogram.

   The image classification framework for this dissertation was adapted from the work of

Vasantha and Bharathi (2011), which used classification of mammogram images using

hybrid features. The researchers work used the Machine Learning package WEKA to

train their dataset using a decision-tree method (Vasantha & Bharathi, 2011). WEKA is a

collection of machine-learning algorithms for data-mining tasks. The classifiers in

WEKA are methods for predicting group membership based on numeric values. In the

Vasantha and Bharathi's research, the researchers used WEKA to create a decision tree

training set using 28 features that were relevant at each node of the tree (see Table 2).

Table 2

Feature Types (Vasantha & Bharathi, 2011)

Feature	Feature
1	Autocorrelation
2	Contrast
3	Homogeneity
4	Correlation 1
5	Correlation 2
6	Cluster Prominence
7	Cluster Shade
8	Dissimilarity
9	Energy
10	Entropy
11	Homogeneity 1
12	Homogeneity 2
13	Maximum Probability
14	Sum Average
15	Sum Variance
16	Sum Entropy
17	Difference Variance
18	Difference Entropy
19	Information Measure of Correlation 1
20	Information Measure of Correlation 2
21	Inverse Difference Normalized
22	Inverse Difference Moment
23	Mean
24	Variance
25	Skewness
26	Kurtosis
27	Energy
28	Entropy

The 25 features used in this dissertation comprise a feature set that is referred to as

Gray Level Co-Occurrence Matrix (GLCM), which has become the standard for medical

image analysis (Gonzalez & Woods, 2002). The GLCM features statistical methods that

consider the spatial relationship; GLCM is also known as the gray-level spatial dependence matrix. Matlab (2012) has functions for obtaining GLCM from images by calculating how often a specific gray-level intensity value occurs. Matlab (2012) file exchange has some examples for GLCM that were used to define most of the requirements for the 25 features types used in the decision tree.

In this dissertation, the researcher identified 21 features that were obtained from the Matlab function texture features. The researcher obtained the remaining four feature types, mean, variance, skewness, and kurtosis, from the state of the histogram-based features. These features are listed in Table 3 and the set of tables in Image Classification Results, where Malignant referred to cancerous cells that had the ability to spread to other areas within the body, and benign referred to an image that had no cancer found.

When classifying an object, one needs to look for the ROI in the image (Vasantha & Bharathi, 2011). The extraction of the ROI, determined by the given thresholding method, is necessary to analyze the area of interest, and to find the prominent features that are representative of the classes of the images. The researcher used the ROI for classification framework analysis categorized using Matlab GLCM and Histogram Image Feature Types [see Figure 40(b) Temporal Subtraction Image ROI Extraction Patient A]. The feature-type sum of squares variance, which is a mathematical approach to determine the dispersal of data points, was added in place of sum variance. The researcher used the Matlab "imtool" function to extract the Malignant areas in images. Overall, the WEKA training set used 30 images containing malignancies that were comprised of 10 temporal-subtraction mammogram images, 10 dual-energy subtraction images, and 10 DDSM mammographic images.

Table 3

Matlab GLCM and Histogram Image Feature Types

| Feature	Feature Type	GLCM	Histogram
1	Autocorrelation	autoc	
2	Contrast	contr	
3	Homogeneity	homop	
4	Correlation 1	corrm	
5	Cluster Prominence	cprom	
6	Cluster Shade	cshad	
7	Dissimilarity	dissi	
8	Energy	energ	
9	Entropy	entro	
10	Homogeneity 1	homon	
11	Sum of Square: Variance	sosvh	
12	Maximum Probability	maxpr	
13	Sum Average	savgh	
14	Sum Variance	svarh	
15	Sum Entropy	senth	
16	Difference Variance	dvarh	
17	Difference Entropy	denth	
18	Correlation 1	inf1h	
19	Correlation 2	inf2h	
20	Inverse Normalized	indnc	
21	Inverse Moment	idmnc	
22	Mean		Mean
23	Variance		Variance
24	Skewness		Skewness
25	Kurtosis		Kurtosis

  (a)   (b)
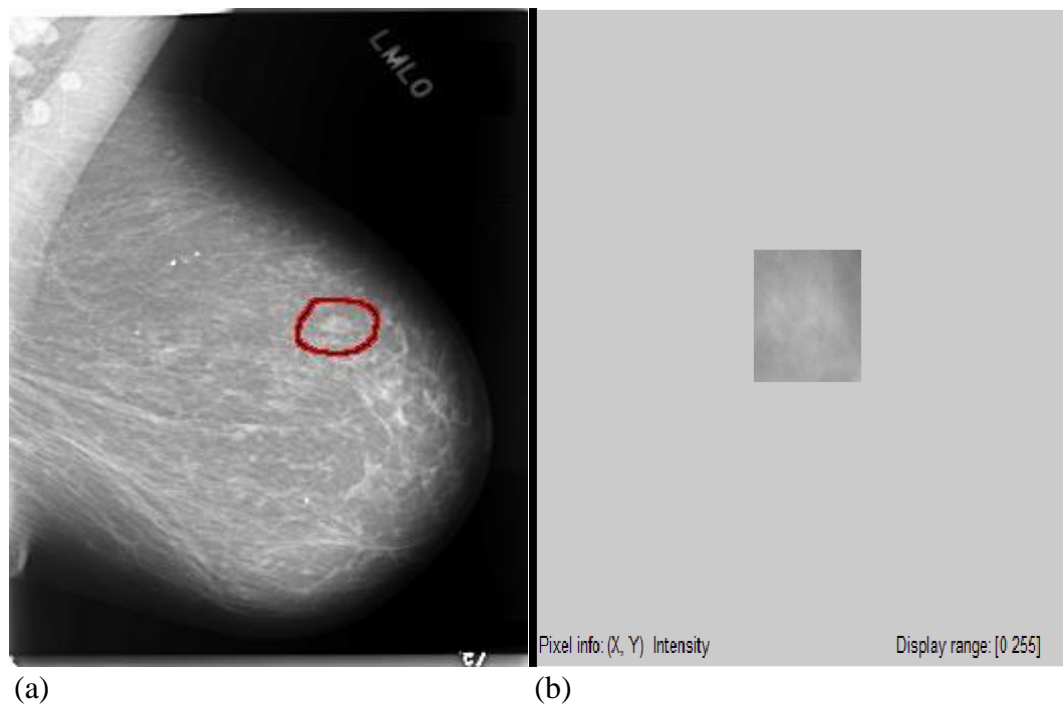
Figure 40. Temporal subtraction image ROI extraction, Patient A.

Image classification results contain the WEKA results from tests completed on accuracy measures of classification by using the features listed in Table 3. Figure 40, Temporal Subtraction Image ROI Extraction Patient A, shows the ROI for the cancer. Figure 40(b), WEKA Precision, Recall, F-measure, shows the results from running Matlab GLCM Image Feature Types outputs. The following summary shows the accuracy measures of classification by using the features selected from temporal subtraction for Patient A. The summary shows the results of precision, recall, and f-measure on DDSM malignant cancer images, which were used in Appendix A to compare the results of the different thresholding methods on temporal subtraction mammogram, Appendix B dual-energy, and Appendix A DDSM mammogram images with malignant cancer.

Precision is the proportion of correctly predicted classes:

$$\text{Precision} = \frac{tp}{tp + fp}$$

Recall is the proportion of the correctly predicted class of true classes:

$$\text{Recall} = \frac{tp}{tp + fn}$$

F-measure is the weighted harmonic mean of precision and recall:

$$\text{F-measure} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

Summary

In Chapter 2, the researcher discussed several hypotheses and questions in accordance with this work, and this chapter supports the exploration of those questions. Data for voxel-matching for temporal subtraction within the registration process was examined using Matlab R2013b, which supports both CT and MRI mammogram images. The results of these CT and MRI mammogram images were the same. The effects of temporal over dual-energy methods are most evident in the global thresholding method where the ground truth was markedly different. One of the positive aspects in working with both temporal and dual-energy images was the high resolution. The predominant negative aspect in working with dual-energy images was that their data were presented with RGB grayscale values (0–255), which must be converted to the standard grayscale images value (0–255) used in temporal images.

Temporal thresholding and dual-energy did hold the ROI, and the ground truth was different between temporal and dual-energy. In global thresholding, neither temporal nor dual-energy methods held the ROI. In local thresholding, the temporal method did hold

the ROI; dual-energy did as well, although dual-energy local was very block-oriented. In

connected-component labeling, neither temporal nor dual-energy held the ROI.

The researcher utilized histograms to find features mean, variance, skewness, and

kurtosis. One of the images using K-means had very good results finding ROI in

temporal subtraction at threshold level .636719, and dual-energy at threshold level

.441176. Regional growing had good results as well, with temporal method using the 8-

neighbors function to return the coordinates of the ROI and dual-energy method using the

4-neighbors function to return the coordinates of the ROI. The split-and-merge technique

did not hold the ROI with temporal or dual-energy methods. It was possible to extract

the objects from the background ground-truth using all thresholding methods, but the real

test was to hold the ROI for malignant growth.

The dissertation researcher did consider thresholding method comparisons in finding

malignant growth in CT and MRI mammogram images, and Matlab handled both

formats. Thresholding methods identified ROI in the original DDSM classification, and

the size of the malignant growth did seem to affect the quality of the result in the

thresholding and classification methods used.

The researcher used classification for measurements that assessed the effectiveness of

the proposed methods. Matlab had some examples of precision, recall, and f-measure for

images, which the researcher used to find 21 texture feature types. The remaining four

features were found from the histogram analysis. Classification results looked at two

areas: malignancy, or cancerous cells, and benign, or an image where no cancer was

found. The researcher utilized WEKA 3.6.10 for classification, and the training set was

derived from a collection of all 30 images containing malignancies. These were

comprised of 10 DDSM mammographic images, 10 dual-energy subtraction images, and 10 temporal subtraction mammogram images. Each image used the two parts of classification: malignancy and benign. Image classification used a J48 decision tree on 25 features types, which the researcher identified and obtained from the Matlab function texture features.

From the results collected, the researcher looked at precision, recall, and f-measure, shown in Appendices A. All values are derived from the experiments described in Chapter 4 using Image classification with WEKA 3.6.10, which yielded precision, recall, and f-measure by running the feature types on each of the two image classifications. The final step was to run the derived test set by applying a thresholding method to a single image against the training set. Each image type has its own decision tree that provided the precision, recall, and f-measure values. The results given are expressed as *precision,* which is the probability of being correct in finding the result of malignancy using the appropriate image classification. *Recall* is the probability of correctly identifying the class, and *F-measure* is the weighted harmonic mean of precision and recall. Appendices A and B contain the values the researcher examined to determine the best combination that yielded the greatest calculated precision values.

Chapter 5

Conclusions, Implications, Recommendations, and Summary

Conclusions

The goal of this dissertation was to create a computer aided diagnosis (CAD) system as a tool to help in the detection of breast cancer in computed tomography (CT) mammography images, in which previous CT mammography images were used to enhance the next series of images. In the first stage of the dissertation, the researcher used image subtraction of images from the same patient over time. This dissertation involved image enhancement by applying image registration and subtraction using temporal image registration and dual-energy subtraction. This dissertation researcher investigated both techniques of subtraction, utilizing Matlab registration to apply temporal subtraction, and employing Lewin's (2003) dual-energy subtraction images for the subtraction section.

The second stage of the experiments involved employing eight different thresholding techniques. The data included 30 images containing malignancies that were composed of 10 DDSM mammographic images, 10 dual-energy subtraction images, and 10 temporal subtraction mammogram images. The first thresholding technique employed was the amplitude thresholding method. When using amplitude thresholding, the researcher discovered that the most noticeable difference between temporal subtraction images and dual-energy images was the ground truth, in which temporal subtraction was white and the dual-energy subtraction was black. The researcher employed the second and third

thresholding techniques using global and local thresholding where the global technique used the Otsu (1979) method, and the local version segmented gray-level images into regions and then thresholded each region separately.

The global thresholding method had high precision, recall, and f-measure, which was misleading because applying the global thresholding method made the entire image fill in with white, in effect comparing two nearly identical data sets. In local thresholding, the dual-energy subtraction method worked the best as cancerous tissue lesions were still visible in the images. The fourth technique used connected-component analysis and did not seem to highlight cancerous tissue lesions as the images became very grainy. The fifth used the overall histogram analysis to show minima and maxima of the image as seen in Chapter 2. As described in Chapter 4, the researcher used the histogram to obtain the mean, variance, skewness, and kurtosis features for image classification. The sixth technique used a K-means at different levels. K-means clustering behaved in an intuitive manner and a split process was expected. It is evident in the Chapter 4 figures that the last images did reach ultimate results because of small, slightly deviant regions. Thresholding with K-mean clustering did, however seem to generally retain the cancerous tissue lesions using temporal subtraction at a threshold level of .636719 and dual-energy subtraction at a threshold level of .441176. The seventh technique used a region-based method called regional growing. This method seemed to work well with both temporal subtraction at the 8-neighbors function and dual-energy subtraction at the 4-neighbors function as the cancerous tissue was noticeable in both methods. The final thresholding method used the split-and-merge technique. The split image was enhanced when viewed in color, while the grayscale contained unrecognizable fragments. Both

temporal subtraction and dual-energy subtraction images did not reveal any noticeable cancerous tissue lesions. Out of the original eight thresholding methods, four techniques were found to be acceptable for the next stage of image classification. The researcher utilized histogram analysis to show minima and maxima of the image, and provided four feature types—mean, variance, skewness, and kurtosis.

A decision tree represented segmentation of data created by applying a series of simple rules. A J48 decision tree method with a classifier base was used to find candidates for cancer by detecting features based on shape and density of the cancerous tissue in the mammography images. The researcher used WEKA 3.6.10 with the J48 decision tree method, and 21 features were identified and obtained from the Matlab function texture features. An additional four feature types were obtained from the state of the histogram-based features. Extraction of the ROI was necessary to analyze the area of interest and to find the prominent features. The classification result malignancy referred to cancerous cells that had the ability to spread to other areas within the body. Finally, the benign classification referred to an image that had no cancer found.

Implications

The effect of this study on future research is that CAD systems using a multiple three-stage approach can help radiologists find cancerous tissue in CT and MRI mammography images using image registration and subtraction, thresholding, and classification for the detection of cancerous tissue. One main result of this research shows how image subtraction and thresholding methods together can enhance mammography images.

The effect of this dissertation on the medical practice will help radiologists and physicians detect breast cancer by using CAD. A patient takes CT mammogram every

year—sometimes every 6 months for high-risk patients for routine checkups; radiologists routinely screen hundreds of these mammogram images for cancerous tissues. The implications of this work could provide a second look that would help overworked radiologists and doctors more effectively read mammogram images using CAD.

Recommendations

The results of this dissertation can be improved further by increasing the size of the study including additional images of DDSM, dual-energy, and temporal images. Additional classification methods also should be examined. Future researchers should consider variations on the thresholding methods, for example, by embedding expert knowledge in the CAD design using thresholding software. Further work can be directed at higher-resolution mammograms. Segmentation techniques may be eliminated through the use of threshold interpolation between regions. Also, a three-dimensional surface plot would present smooth, readily interpreted surfaces.

Future researchers can pursue several areas as an extension of this research. First, increasing the size of the study by increasing the number of images used may produce more thorough results. A larger study yields greater statistical significance. Additional specialized studies are possible, because of different factors such as aging, body mass index (BMI), size, number of pregnancies, and breast augmentation along with many other factors. Another extension of this research would employ different classification methods such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Genetic Algorithms (GA), Fuzzy Support Vector Machines (FSVM), and Genetic Algorithms with Neural Networks, all of which have been used for image classification.

The extension of using multiple thresholding methods, or different combinations of

the methods presented here, may provide better overall results. The researcher combined temporal with amplitude thresholding, which held the highest precision, recall, and f-measure, along with dual-energy with local thresholding, which held the second highest thresholding for precision, recall, and f-measure results. Local thresholding broke down the original image into smaller sub-regions, and each sub-region was given a threshold individually. Amplitude thresholding could be then applied at this step at the smaller sub-regions levels. Combining the use of temporal-based and dual-energy enhanced digital mammography from a patient into a single technique may provide the benefits of both approaches.

Summary

The goal of this research was to assess the effectiveness of various thresholding methods in the context of a three-stage approach, to help radiologists find cancerous tissue lesions in CT and MRI mammography images. This three-stage approach included image registration and subtraction, thresholding, and classification for the detection of cancerous tissue with CAD system design. The researcher identified two independent variables: subtraction method in stage one (temporal subtraction and dual-energy subtraction), and thresholding method in stage two. The remaining factors were fixed (including the classification method). All the experiments used the Matlab's 2012a Environment to obtain relevant measurements and WEKA 3.6.10 for image classification. Amplitude and K-means thresholding scored high in image classification with dual-energy subtraction, but the local thresholding method provided the best precision, recall, and f-measure results.

The goal of this dissertation was to design an application to help radiologists and

physicians better detect breast cancer. The three-stage approach of image registration and subtraction, application of thresholding techniques, and classification detailed within this dissertation achieved this goal. While additional research can enhance the process, the methods presented here are a solid foundation for providing automated assistance to radiologists and physicians in detecting and diagnosing suspected breast cancer.

Appendix A

Image Results

The data values in Appendix A are derived from the experiments detailed in Chapter 4 and were drawn from three different image types. All values are expressed in milliseconds. The Matlab profiling tool MuPAD® process is used to isolate time spent thresholding on time() which returns the total CPU time in milliseconds.
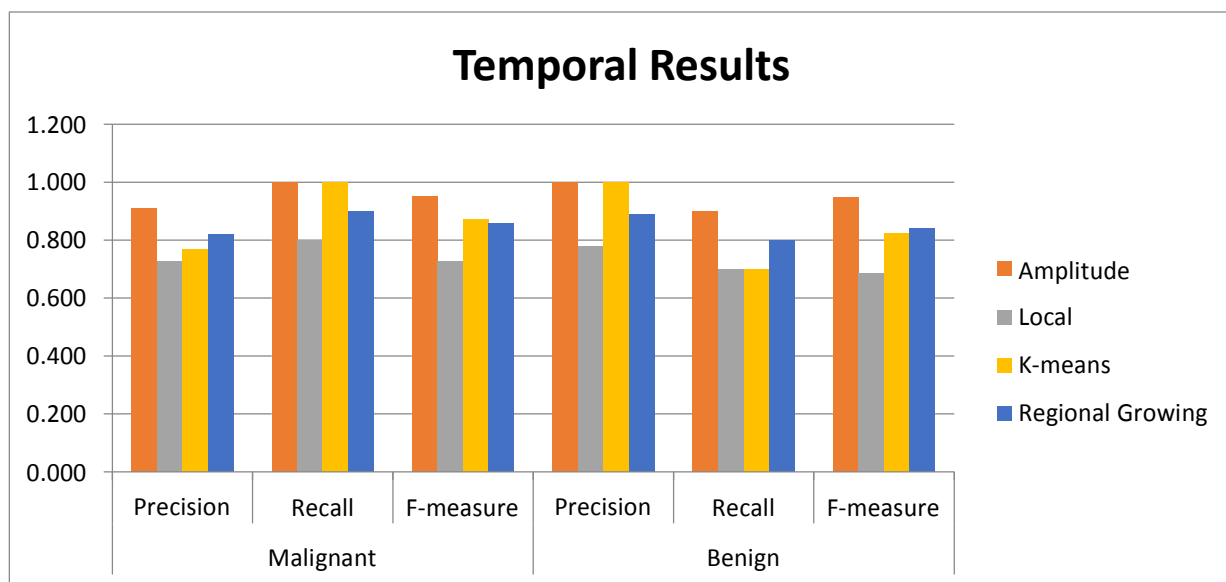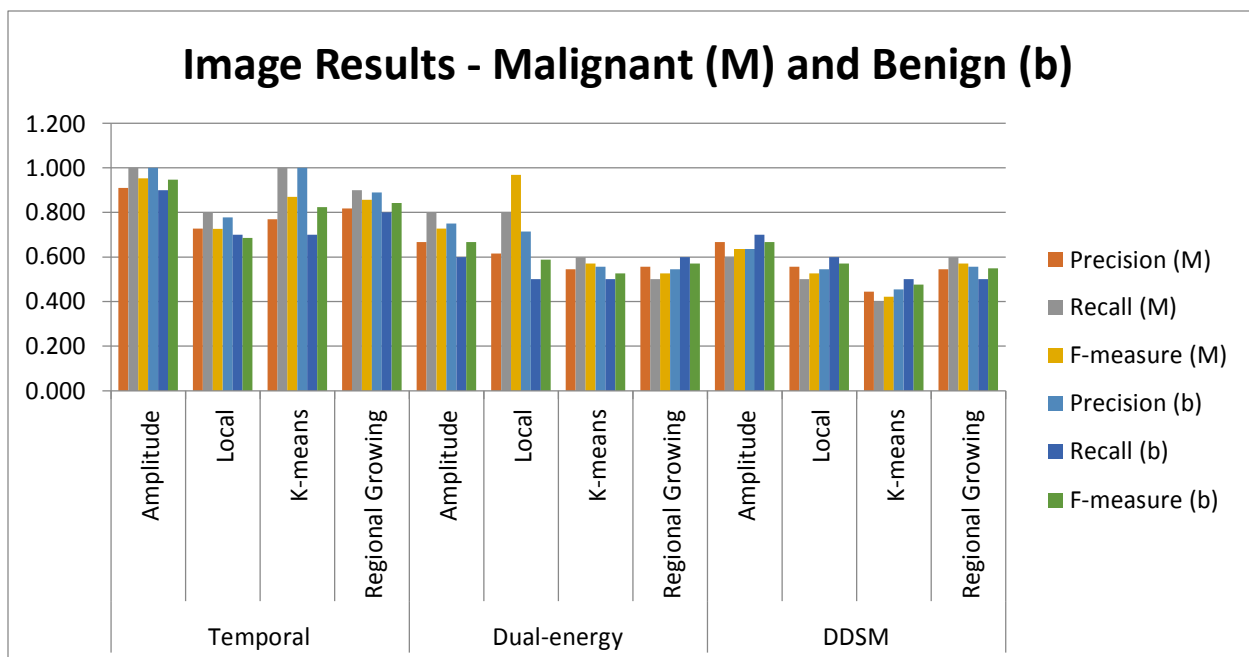
Appendix A data relate to temporal subtraction images, dual-energy subtraction images, and Digital Database for Screening Mammography (DDSM) images. There are 10 unique samples from each image type providing a total collection of 30 images. The methodology used to calculate precision, recall, f-measure, and CPU timing utilized averages over multiple images for each of the four threshold methods in a three stage approach:
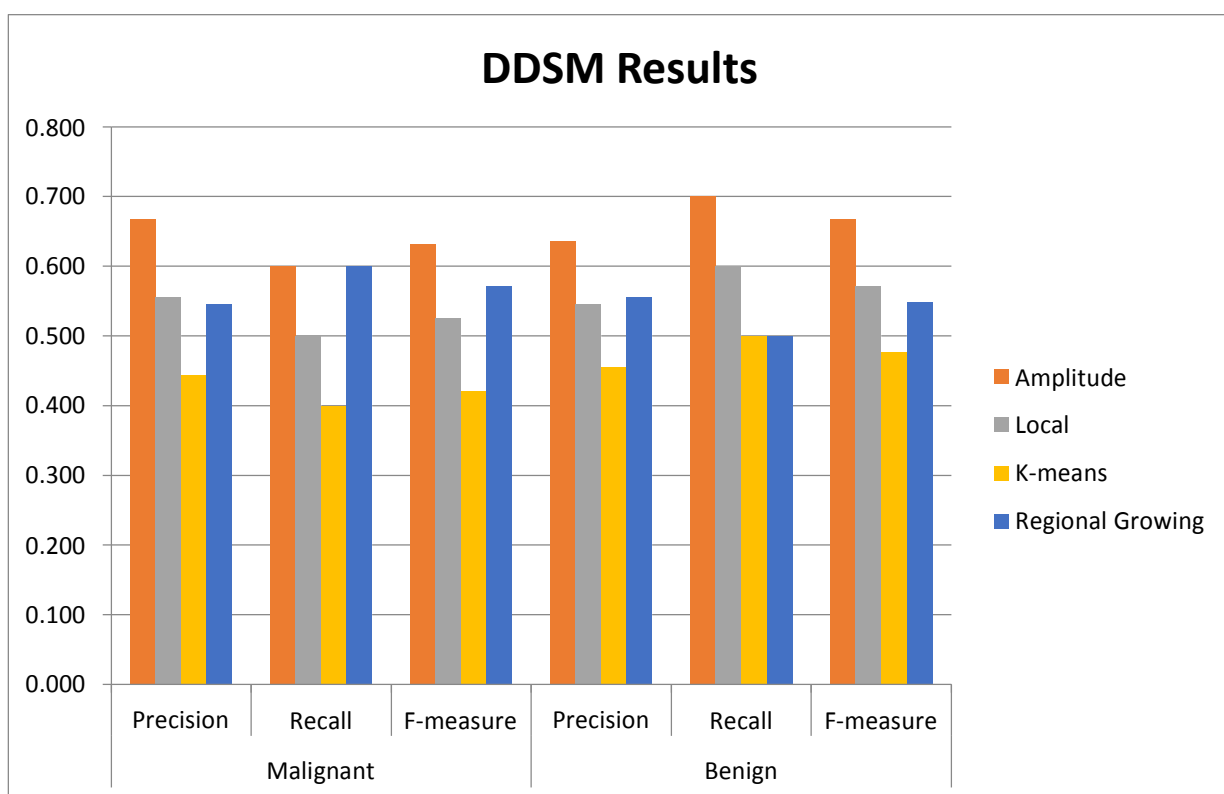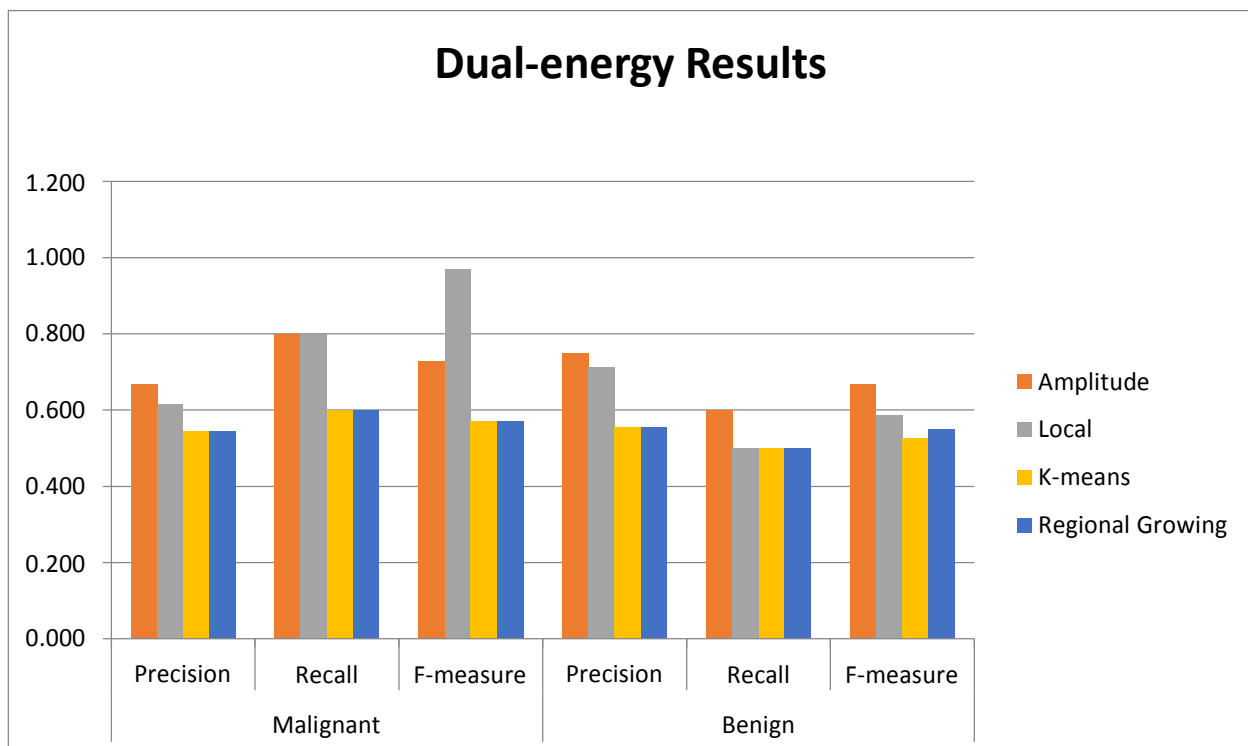
1. Image registration – Defines each image type (temporal subtraction, dual-energy subtraction, and DDSM) by determining regions within each image. A region is classified as *malignant* or *benign*.

2. Thresholding – Provides the test set for each image type (temporal subtraction, dual-energy subtraction, and DDSM) using four different thresholding techniques—amplitude, local, k-means, and regional growing. Histogram analysis provided four feature types—mean, variance, skewness, and kurtosis. Matlab 2012a grey-level, co-occurrence matrix (GLCM) provided 21 target features types which the researcher obtained from the Matlab function texture features. After thresholding, the 25 texture feature type values were computed for the *malignant* and *benign* area of each image.

3. Image classification – Combines the test set and training set from thresholding

to produce the precision, recall, and f-measure. An image classification

decision tree calculated the final values from WEKA 3.6.10 along with

WEKA J48 classifier and cross-validation 10 fold.

The following table contains the raw values for precision, recall, and f-measure for each threshold method by image type for images containing malignant and benign artifacts. It is followed first by a chart depicting all image types combined and then one for each of the individual image types.

Image type	Threshold Method	Malignant			Benign		
		Precision (M)	Recall (M)	F-measure (M)	Precision (b)	Recall (b)	F-measure (b)
Temporal	Amplitude	0.909	1.000	0.952	1.000	0.900	0.947
	Local	0.727	0.800	0.726	0.778	0.700	0.685
	K-means	0.769	1.000	0.870	1.000	0.700	0.824
	Regional Growing	0.818	0.900	0.857	0.889	0.800	0.842
Dual-energy	Amplitude	0.667	0.800	0.727	0.750	0.600	0.667
	Local	0.615	0.800	0.969	0.714	0.500	0.588
	K-means	0.545	0.600	0.571	0.556	0.500	0.526
	Regional Growing	0.556	0.500	0.526	0.545	0.600	0.571
DDSM	Amplitude	0.667	0.600	0.632	0.636	0.700	0.667
	Local	0.556	0.500	0.526	0.545	0.600	0.571
	K-means	0.444	0.400	0.421	0.455	0.500	0.476
	Regional Growing	0.545	0.600	0.571	0.556	0.500	0.549

Image Results - Malignant (M) and Benign (b)

Temporal Results
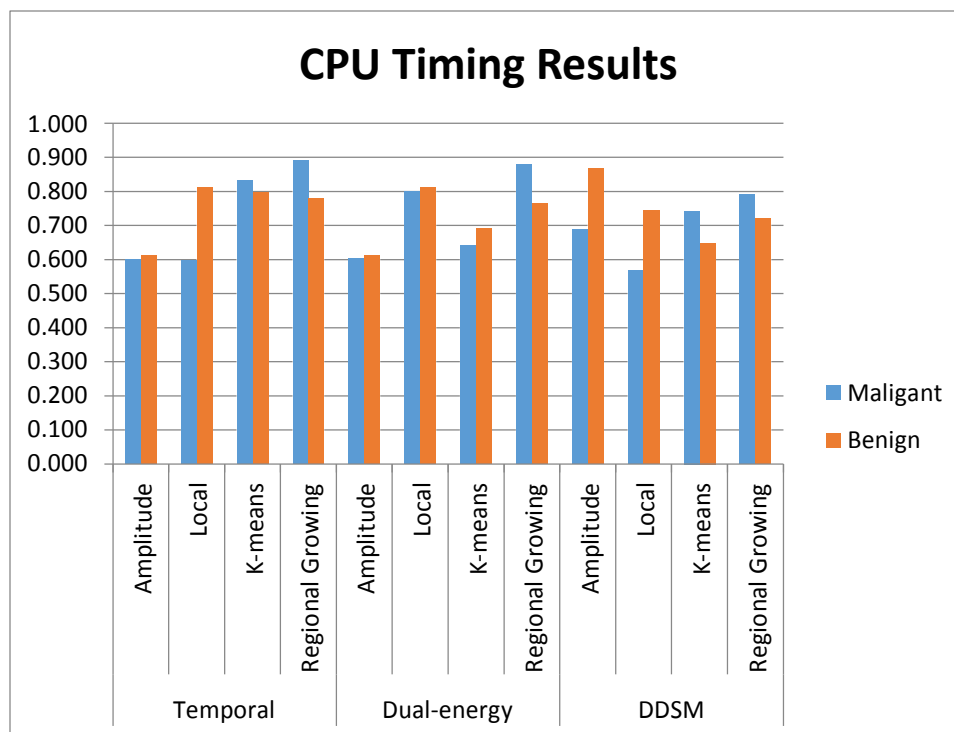
Dual-energy Results

DDSM Results

Appendix B

Runtime Analysis

Appendix B contains the runtime analysis framework on experimental timing (averaged over multiple images) that the researcher computed for each of the four threshold methods and reported as average CPU time in milliseconds. The raw values are presented first and are followed by a graph depicting the results in each category.

Image type	Threshold Method	Maligant CPU Timing	Benign CPU Timing
Temporal	Amplitude	0.600	0.611
	Local	0.598	0.812
	K-means	0.832	0.798
	Regional Growing	0.892	0.780
Dual-energy	Amplitude	0.604	0.611
	Local	0.800	0.812
	K-means	0.640	0.690
	Regional Growing	0.879	0.765
DDSM	Amplitude	0.689	0.867
	Local	0.567	0.744
	K-means	0.742	0.649
	Regional Growing	0.792	0.722

CPU Timing Results

References

Al-Hyari, A., Al-Taee, A., & Al-Taee, M. (2013). Clinical decision support system for diagnosis and management of chronic renal failure. *IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies*, *Amman, Jordan*. doi:10.1109/AEECT.2013.6716440

Arora, R., & Suman. (2012). Comparative analysis of classification algorithms on different datasets using WEKA. *International Journal of Computer Applications, 54*(13), 21–25. http://dx.doi.org/10.5120/8626-2492

Ballard, D. H., & Brown, C. M. (1982). Computer vision. Englewood Cliffs, NJ: Prentice Hall.

Brink, A. (1991). Comments on grey-level thresholding of images using a correlation criterion. *Pattern Recognition Letters, 12*, 91–92. doi:10.1016/0167-8655(91)90054-P

Burt, P. J., & Adelson, E. H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications, 31*(4), 532–540. http://dx.doi.org/10.1109/tcom.1983.1095851

Carton, A-K., Lindman, K., Ullberg, C., Francke, T., & Maidment, A. D. A. (2007). Dual-energy subtraction for contrast-enhance digital breast tomosynthesis. *Proceedings of Medical Imaging 2007: Physics of Medical Imaging,* USA, 651007. doi:10.1117/12.713703

Chen, Y., & Chen, O. (2006). Robust image segmentation using modified edge-following scheme with automatically-determined thresholds. *First International Conference on Innovative Computing, Information, and Control, 3,* 292–295.

http://dx.doi.org/10.1109/icicic.2006.511

Chow, C. K., & Kaneko, T. (1972). Automatic boundary detection of the left ventricle
from cineangiograms. *Computers and Biomedical Research, 5*, 388–410.
http://dx.doi.org/10.1016/0010-4809(72)90070-5

Csetverikov, D. (2012). *Basic algorithms for digital image analysis*. Retrieved from
http://progmat.uw.hu/oktseg/kepelemzes/lec09_threshold_4.pdf

Dembczynski, K., Waegeman, W., Cheng, W., & Hullermeier, E. (2011). An exact
algorithm for f-measure maximization. *Neural Information Processing Systems*.
Retrieved from http://www.weiweicheng.com/research/slidesposters/cheng-
nips11poster.pdf

Dromain, C., Thibault, F., Diekmann, F., Fallenberg, E. M., Jong, R. A., Koomen, M., . .
. Toledano, A. (2012). Dual-energy contrast-enhanced digital mammography:
Initial clinical results of a multireader, multicase study. *Breast Cancer Research,
14*, R94. http://dx.doi.org/10.1186/bcr3210

Eltonsy, N. H., Tourassi, G. D., & Elmaghraby, A. S. (2007). A concentric morphology
model for the detection of masses in mammography. *IEEE Transactions on
Medical Imaging, 26*, 880–889. http://dx.doi.org/10.1109/tmi.2007.895460

Gilbert, F. J., Astley, S. M., Gillian, M. G., Agbaje, O. F., Wallis, M. G., James, J., . . .
Duffy, S. W. (2008). Single reading with computer-aided detection for screening
mammography. *New England Journal of Medicine, 359,* 1675–1684.
doi:10.1056/NEJMoa0803545

Gonzalez, R. C., & Woods, R. E. (2002). *Digital image processing* (2nd ed.). Reading,
MA: Addison-Wesley.

Gonzalez, R. C., Woods, R. E., & Eddins, S. L. (2012). *Image processing place*.
Retrieved from http://www.imageprocessingplace.com/index.htm

Gupta, D. L., Malviya, A. K., & Singh, S. (2012). Performance analysis of classification
tree learning algorithms. *International Journal of Computer Applications, 55*, 39–
44. http://dx.doi.org/10.5120/8762-2680

Hornak, J. P. (1999). *The basics of MRI*. Retrieved from
https://www.cis.rit.edu/htbooks/mri/inside.htm

Hounsfield, G. N. (1973). Computerized transverse axial scanning (tomography): Part I.
Description of system. *British Journal of Radiology, 46*, 1016–1022. Retrieved
from http://www.ncbi.nlm.nih.gov/pubmed/4757352

Jain, A. K. (2002). *Fundamentals of digital image processing*. Upper Saddle River, NJ:
Prentice Hall.

Itai, Y., Kim, H., Ishikawa, S., Katsuragawa, S., & Doi, K. (2008). A new registration
method with voxel-matching technique for temporal subtraction images.
Proceedings of Medical Imaging 2008: Computer-Aided Diagnosis, USA,
69153I. doi:10.1117/12.769921

Karellas, A., Vedantham, S., & Lewin, J. (2009). Digital mammography: From planar
imaging to tomosynthesis. *Medicamundi, 53*(1), 14–19. Retrieved from
http://incenter.medical.philips.com/doclib/enc/fetch/2000/4504/577242/577256/5
88821/5050628/5313460/5983378/05_Karellas.pdf%3fnodeid%3d5983979%26v
ernum%3d1

Kroon, D. (2008). *Region growing*. Retrieved from
http://www.mathworks.com/matlabcentral/fileexchange/19084

Kulkarni, M., & Nicolls, F. (2009). *Interactive image segmentation using graph cuts*. Retrieved from http://www.dip.ee.uct.ac.za/~nicolls/publish/mk09-prasa.pdf

Kunt, M., Ikonomopoulos, A., & Kocher, M. (1985). Second-generation image-coding techniques. *IEEE, 73,* 549–574. http://dx.doi.org/10.1109/proc.1985.13184

Lewin, J. L. (2003). Dual-energy contrast-enhanced digital subtraction mammography feasibility. *Radiology, 229*(1). Retrieved from http://radiology.rsna.org/content/229/1/261.full

Li, Z., & Zhang, J. (2006). Image segmentation based on inscribed circle. *The 18th International Conference on Pattern Recognition, 2*, 247–250. http://dx.doi.org/10.1109/icpr.2006.689

Lin, Z., Jin, J., & Talbot, H. (2001). Unseeded region growing for 3-d image segmentation. *Visualization 2000 Pan-Sydney Workshop and Practice, 2*. Retrieved from http://crpit.com/confpapers/CRPITV2Lin.pdf

Linh, D. L., & Linh, H. Q. (2010). *Medical image registration in Matlab*. Retrieved from http://www.fas.hcmut.edu.vn/webhn10/Baocao/PDF/TDLinh-MIRegistration.pdf

Matlab. (2012). *Matlab and Simulink home page* [Database]. Retrieved from http://www.mathworks.com/

Matlab. (2013). *Measure distance between pixels in Image Viewer*. Retrieved from http://www.mathworks.com/help/images/measuring-the-distance-between-two-pixels.html

Maurer, C. R. Jr., & Fitzpatrick, J. M. (1993). A review of medical image registration. In R. J. Maciunas (Ed.), *Interactive image-guided neurosurgery* (17–44). Rolling Meadows, IL: American Association of Neurological Surgeons.

Miyake, N., Kim, Itai, & Ishikawa. (2009). Automatic detection of lung nodules in temporal subtraction image by use of shape and density features. *2009 Fourth International Conference on Innovative Computing, Information and Control*, 1288–1292. http://dx.doi.org/10.1109/icicic.2009.118

Nock, R., & Nielsen, F. (2004). Statistical region merging. *IEEE Transaction on Pattern Analysis and Machine Intelligence, 26*(11), 1452–1458. http://dx.doi.org/10.1109/tpami.2004.110

Olson, D., & Delen, D. (2008). *Advanced data mining techniques*. New York, NY: Springer.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics Society, 9*(1), 62–66. http://dx.doi.org/10.1109/tsmc.1979.4310076

Pal, M. R., & Pal, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition, 26*, 1277–1294. http://dx.doi.org/10.1016/0031-3203(93)90135-j

Pavlidis, T. (1980). *Structural pattern recognition*. New York, NY: Springer-Verlag.

Pavlidis, T. (1982). *Algorithms for graphics and image processing*. Rockville, MD: Computer Science Press.

Prewitt, J. M. S., & Mendelsohn, M. J. (1966). The analyst of cell images. *Annual of the New York Academy of Sciences*, *128*, 1035–1053. http://dx.doi.org/10.1111/j.1749-6632.1965.tb11715.x

Rafferty, E. A. (2007). Digital mammography: Novel applications. *Radiologic Clinics of North America, 45*(1), 831–843. http://dx.doi.org/10.1016/j.rcl.2007.06.005

Rosenfeld, A., & Kak, A. C. (1982). *Digital picture processing*. Orlando, FL: Academic

Press.

Russell, S., & Norvig, P. (2003). *Artificial intelligence: A modern approach*. Upper

Saddle River, NJ: Prentice Hall.

Saidin, N., Ngah, U. K., Sakim, H. A. M. S., Siong, D. N., Hoe, M. K., & Shuaib, I. L.

(2010). Density based breast segmentation for mammograms using graph cut and

seed based region growing techniques. *2010 Second International Conference on

Computer Research and Development* (pp. 246–250).

http://dx.doi.org/10.1109/iccrd.2010.87

Salma, G. I., Abdelhalim, M. B., & Zeid, M. A. (2013). Breast cancer diagnosis on three

different datasets using multi-classifiers. *International Journal of Computer and

Information Technology, 1*, 36–43. Retrieved from

http://ijcit.com/archives/volume1/issue1/Paper010105.pdf

Sampat, M. P., Markey, M. K., & Bovik, A. C. (2005). Computer-aided detection and

diagnosis in mammography. In A. Bovik (Ed.) *Handbook of Image and Video

Processing* (pp. 1195–1217). http://dx.doi.org/10.1016/b978-012119792-6/50130-

3

Singh, S., & Al-Mansoori. (2000). Identification of regions of interest in digital

mammograms. *Journal of Intelligent Systems, 10*, 183–217.

http://dx.doi.org/10.1515/jisys.2000.10.2.183

Tou, J. T., & Gonzalez, R. C. (1974). *Pattern recognition principles*. Reading, MA:

Addison Wesley.

Tourassi, G. D., Vargas-Voracek, R., & Floyd, C. E. (2003). Content-based image

retrieval as a computer aid for the detection breast cancer. *Proceedings of SPIE*

5032, Medical Imaging 2003, University of South Florida. doi:10.1117/12.481105

Vasantha, M., & Bharathi, V. S. (2011). Classifications of mammogram images using hybrid features. *European Journal of Scientific Research, 57*(1), 87–96. Retrieved from http://www.eurojournals.com/ejsr.htm

Verbeek, P. W., Vrooman, H. A., & Vliet, V. (1988). Low-level image processing by max-min filters. *IEEE Signal Processing, 15*(3), 249–258. http://dx.doi.org/10.1016/0165-1684(88)90015-1

White, R. (1990). The learning rate in back-propagation systems: An application of Newton's method. *International Joint Conference on Neural Networks,* 1, 679–684. http://dx.doi.org/10.1109/ijcnn.1990.137647

Wu, X. (1992). Image coding by adaptive tree-structured segmentation. *IEEE Transactions on Information Theory, 38*(6), 1755–1767. http://dx.doi.org/10.1109/18.165448

Xiong, G. (2005). Local adaptive thresholding. Retrieved from http://www.mathworks.com/matlabcentral/fileexchange/authors/19310

Zitová, B., & Flusser, J. (2003). Image registration methods: A survey. *Image Vision Computing, 21,* 977–1000. http://dx.doi.org/10.1016/s0262-8856(03)00137-9