ENHANCING THE IMAGE QUALITY OF DIGITAL BREAST TOMOSYNTHESIS

A Dissertation
Presented to
The Academic Faculty

By

Steve Si Jia Feng

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Biomedical Engineering

Georgia Institute of Technology and Emory University

AUGUST 2013

ENHANCING THE IMAGE QUALITY OF DIGITAL BREAST TOMOSYNTHESIS

Approved by:

Dr. Ioannis Sechopoulos, PhD Advisor
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Srini Tridandapani, MD, PhD
Department of Radiology
Emory University

Dr. John Oshinski, PhD
Department of Biomedical Engineering
Georgia Institute of Technology

Dr. Carl D'Orsi, MD
Department of Radiology
Emory University

Dr. Xiangyang Tang, PhD
Department of Bioengineering
Georgia Institute of Technology

Date Approved: 4/30/2013

Dedicated to my parents, Henry C. Fung and Louisa K. Fung, who have always supported me in each and every endeavor I sought.

ACKNOWLEDGEMENTS

This dissertation is the culmination of years of work and the contributions and efforts of many people. To begin with, my education, development, and growth as a scientist, as an engineer, and as a person were aided and abetted in more ways than I can count by Dr. Ioannis Sechopoulos. He was always there to provide advice when I had grown tired of beating my head against the same roadblock, to provide assistance when I took a wrong turn and got lost, and to provide support when I just needed an extra hand pushing the stalled car that is my dissertation up a steep hill. And last but not least, he was there to extend a hand to help me up when I fell. His intellect, drive, and mentorship are unparalleled and it saddens me to think that I will never be able to repay his efforts.

I would also like to thank Dr. John Oshinski, Dr. Xiangyang Tang, Dr. Srini Tridandapani, and Dr. Carl D'Orsi for their advice, service, and efforts on my behalf.

A special thanks goes out to Ms. Jessica Paulishen for her efforts coordinating our work, shielding me from bureaucracy, and most importantly of all, ensuring that I never ran out of coffee. Mrs. Ellen D'Orsi, whose steady hand, timely assistance, and kind understanding in the Breast Imaging Center were invaluable, deserves special mention. I would also like to thank Ms. Linda Burr for helping me get settled when I first arrived at Winship.

I would like to extend my gratitude to Dr. James Nagy and Dr. Veronica Bustamante. Their innovations and efforts were invaluable to our collaboration.

And none of this would have been possible without my family's never-ending support and unconditional love.

This work was supported in part by Grant R01CA163746 from the National Cancer Institute, the PHS Grant UL1 RR025008 from the Clinical and Translational

Science Award program, National Institutes of Health, National Center for Research Resources, Grant DMS-1115627 from the National Science Foundation, and a grant from the Atlanta Clinical and Translational Science Institute and University Research Committee, Emory University.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	ix
LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	xiii
LIST OF SYMBOLS	xvi
SUMMARY	xvii
CHAPTER	
1 Introduction	1
1.1 General Background	1
1.2 Digital Breast Tomosynthesis	3
1.3 Approach – Novel Software-Based Solutions	6
2 A Software-Based X-ray Scatter Correction Method	11
2.1 Significance	12
2.2 Materials and Methods	13
2.3 Results	25
2.4 Discussion	36
2.5 Conclusion	39
3 Clinical Digital Breast Tomosynthesis System: Dosimetric Characterization	40
3.1 Significance	40
3.2 Materials and Methods	41
3.3 Results	47
3.4 Discussion	54

3.5 Conclusion 57

4 Objective Models of Compressed Breast Shapes Undergoing Mammography 58

4.1 Significance 59

4.2 Materials and Methods 60

4.3 Results 65

4.4 Discussion 78

4.5 Conclusion 80

5 PCA Model-Based X-Ray Scatter Map Library 81

5.1 Significance 81

5.2 Materials and Methods 82

5.3 Results 91

5.4 Discussion 93

5.5 Conclusion 95

6 Spectral reconstruction algorithm 96

6.1 Significance 96

6.2 Materials and Methods 98

6.3 Results 103

6.4 Discussion 111

6.5 Conclusion 113

7 X-Ray Scatter Map 2DPCA Model 114

7.1 Significance 114

7.2 Materials and Methods 115

7.3 Results 119

7.4 Discussion 124

7.5 Conclusion 125

8 Conclusion 126

REFERENCES 128

VITA 137

LIST OF TABLES

	Page
Table 2.1: Ratio of Maximum Signal Difference (SD)	30
Table 2.2: Ratio of Signal-Difference to Noise Ratio (SDNR)	30
Table 2.3: Ratio of Integrated Mass Signal (IMS)	31
Table 3.1: AEC Selections for DBT	48
Table 3.2: AEC Selections for FFDM	48
Table 3.3: AEC Selections for Phantoms	49
Table 3.4: D_gN_{DBT} MC Results	51
Table 3.5: D_gN_{FFDM} Monte Carlo Results	51
Table 3.6: MGD_{DBT} Calculated Results	52
Table 3.7: MGD_{FFDM} Calculated Results	52
Table 3.8: Ratio of MGD_{DBT} to MGD_{FFDM}	54
Table 4.1: Percentage of total variance	67
Table 4.2: Average Distance Error	68
Table 5.1: Selected Values for Principal Components of Library	86
Table 5.2: Range and Granularity of Principal Components of Library	92
Table 5.3: Observer Study Results	93
Table 6.1: Signal Difference to Noise Ratio of Glandular Masses	105
Table 6.2: X-ray Tube Voltage and Tube Current-Exposure Time Products	108
Table 6.3: Signal Difference to Noise Ratio of Dual Spectrum Images	109
Table 6.4: Radiation Dose Resulting from Dual-Spectrum Imaging	111
Table 7.1: Selected Values for Principal Components of the Breast Shape	116
Table 7.2: Percentage of total variance	120

Table 7.3: 3D Breast Test Volumes 121

LIST OF FIGURES

	Page
Figure 1.1: The geometry of digital breast tomosynthesis imaging	4
Figure 1.2: Typical X-ray spectra used in BT imaging	8
Figure 2.1: The breast tomosynthesis system	16
Figure 2.2: Monte Carlo simulations of the breast in the CC view	17
Figure 2.3: The CIRS Model 082 Breast Mammography Phantom	22
Figure 2.4: Integrated Mass Signal	24
Figure 2.5: Reconstructions of 5 cm thick Phantom Images	26
Figure 2.6: Reconstructions of 8 cm thick Phantom Images	28
Figure 2.7: Offset Gray Level Signal Profiles of Spheroid Masses	32
Figure 2.8: Comparison of MTF before and after noise filtration	33
Figure 2.9: Reconstructions of Patient A Images	34
Figure 2.10: Reconstructions of Patient B Images	35
Figure 3.1: The Hologic Selenia Dimensions system	43
Figure 3.2: X-ray Spectra Models	50
Figure 3.3: Mean Glandular Dose vs Breast Thickness	53
Figure 4.1: Average Distance Error Calculation	64
Figure 4.2: Breast Edges	66
Figure 4.3: Modeled breast shapes using 2 and 6 principal components	69
Figure 4.4: Box-Whisker Plot of Average Distance Error	70
Figure 4.5: PCA Parameter Distributions	71
Figure 4.6: Novel Breast Shapes Generated by the 6-component PCA Model	73
Figure 4.7: Varying PCA parameters of the 6-component CC view model	75

Figure 4.8: Varying PCA parameters of the 6-component MLO view model 76

Figure 4.9: Relationship Between α and Breast Size 77

Figure 5.1: Generation of the 3D Breast Volume from a 2D Breast Shape 85

Figure 5.2: Scatter Correction Observer Study Example Images 90

Figure 6.1: The CIRS Model 082 Breast Mammography Phantom 101

Figure 6.2: Reconstructions of Phantom Images 104

Figure 6.3: Regions of Interest of Reconstructions of Phantom Images 106

Figure 6.4: Regions of Interest of Reconstructions of Patient Images 107

Figure 6.5: Modeled X-ray Spectra 108

Figure 6.6: DBT Reconstructions of the Simulated Dual Spectrum Acquisition 110

Figure 6.7: Zoomed-In Views of the DBT Reconstructions 110

Figure 7.1: Novel Scatter Map of Test Volume A 122

Figure 7.2: Novel Scatter Map of Test Volume B 123

LIST OF ABBREVIATIONS

2D	two dimensional
2DPCA	two dimensional principal component analysis
2DS	2D breast shape
3D	three dimensional
3DV	3D breast volume
ADE	average distance error
AEC	automatic exposure control
CC	cranio-caudal
CEBT	contrast enhanced breast tomosynthesis
CND	chest to nipple distance
CT	computed tomography
DBT	digital breast tomosynthesis
DgN_B	normalized mean glandular dose of Base set
DgN_{DBT}	normalized mean glandular dose resulting from tomosynthesis
DgN_{FFDM}	normalized mean glandular dose resulting from mammography

DgN_{HE}	normalized mean glandular dose of High Energy set
DMIST	Digital Mammographic Imaging Screen Trial
DS	simulated dual-spectrum single-acquisition
FBP	filtered backprojection
FFDM	full field digital mammography
GPU	graphical processing unit
HVL	half value layer
IMS	integrated mass signal
LCW	length along the chest wall
LRF	line response function
MAE	mean absolute error
MC	Monte Carlo
MGD_B	total mean glandular dose of Base set
MGD_{DBT}	total mean glandular dose resulting from tomosynthesis
MGD_{DS}	total mean glandular dose of DS set
MGD_{FFDM}	total normalized mean glandular dose resulting from mammography
MLEM	Maximum Likelihood-Expectation Maximization

MLO	medio-lateral oblique
MTF	modulation transfer function
O	original
PCA	principal component analysis
RGD	relative glandular dose
RMSE	root mean square error
ROI	region of interest
SC	scatter correction/corrected
SD	maximum signal difference
SDNR	signal difference-to-noise ratio
SID	source to imager distance
SIRT	simultaneous iterative reconstruction technique
SPR	scatter-to-primary ratio
SR	spectral reconstruction
Th	compressed breast thickness

LIST OF SYMBOLS

θ	acquisition angle
μ	mean
σ	standard deviation

SUMMARY

A novel imaging technology, digital breast tomosynthesis (DBT), is a technique that overcomes the tissue superposition limitation of conventional mammography by acquiring a limited number of X-ray projections from a narrow angular range, and combining these projections to reconstruct a pseudo-3D image. The emergence of DBT as a potential replacement or adjunct to mammographic screening mandates that solutions be found to two of its major limitations, namely X-ray scatter and mono-energetic reconstruction methods. A multi-faceted software-based approach to enhance the image quality of DBT imaging has the potential to increase the sensitivity and specificity of breast cancer detection and diagnosis. A scatter correction (SC) algorithm and a spectral reconstruction (SR) algorithm are both ready for implementation and clinical evaluation in a DBT system and exhibit the potential to improve image quality. A principal component analysis (PCA) based model of breast shape and a PCA model of X-ray scatter optimize the SC algorithm for the clinical realm. In addition, a comprehensive dosimetric characterization of a FDA approved DBT system has also been performed, and the feasibility of a new dual-spectrum, single-acquisition DBT imaging technique has also been evaluated.

CHAPTER 1

INTRODUCTION

This thesis documents the development of methods aimed at improving the image quality of a novel imaging technology, digital breast tomosynthesis (DBT), as well as understanding its dosimetric characteristics and exploring its potential for dual spectrum imaging. This chapter describes DBT, its potential role in breast cancer imaging, and the challenges DBT faces in becoming a replacement or adjunct to conventional mammography as a screening examination, including the current lack of X-ray scatter reduction methods and the detriments of conventional reconstruction methods.

1.1 General Background

Breast cancer is the second most prevalent cancer in women, behind only skin cancer, in the US. The rate of breast cancer incidence among women in North America is the highest in the world and has increased from one in 20 in 1960 to one in eight in 2006.[1] The American Cancer Society estimates that in 2010, over 207,000 new cases of invasive breast cancer were diagnosed, and that almost 40,000 women died as a result of this disease.[2] This type of cancer is responsible for 3% of all female deaths and is the second leading cause of cancer-related death in women (second only to lung cancer).[3] The 5-year patient survival rate is as high as 98% when breast cancer is diagnosed early and as low as 26% when diagnosis occurs at a late stage after the onset of metastasis.[4] As early diagnosis of breast cancer is the key to a lower mortality rate, annual screening mammography has been recommended for asymptomatic women over the age of 40.[3] During a screening mammogram examination, two X-ray projection images are acquired of each patient's breasts: the cranio-caudal (CC) view in which the breast is compressed horizontally and the medio-lateral oblique (MLO) view in which the breast is compressed

at approximately a seventy degree angle. A radiologist, ideally specializing in breast imaging, interprets, or "reads", all four mammograms searching for abnormal features. These features are separated into inconclusive and benign findings; the former could be associated with breast cancer while the latter is a feature that is not part of the normal breast tissue (e.g. intramammary lymph nodes) but that is not associated with breast cancer.

If the screening examination is normal (no findings or benign findings only), the patient is to return in one year for her next annual screening. The presence of inconclusive findings dictates that additional imaging is required in the form of a diagnostic work-up, which leads to greater patient discomfort and anxiety and higher resulting radiation dose to the patient. The radiologist then studies the screening and diagnostic images to determine if the original inconclusive finding(s) could be indicative of breast cancer or a false alarm. In the case of a false alarm, the patient is sent home with the recommendation that she return for her next annual screening mammogram. However, if the interpreting radiologist determines that the inconclusive finding(s) could actually be breast cancer, the patient is then to undergo a biopsy or a six-month follow-up examination. The sensitivity of screening mammography for breast cancer (proportion of cancer cases identified out of all true cancer cases) is 78.7%, while the specificity for breast cancer (proportion of healthy cases identified out of all true healthy cases) is 89.5% .[5] In addition, the recall rate (proportion of patients subject to a diagnostic work-up) is 10.9% and the positive predictive value of screening mammography is 4.4%.[6] Thus, more than one out of five cancers are missed during screening; while for each patient who actually has breast cancer and who is recalled for the diagnostic work-up, nineteen healthy women are recalled and subjected to what is ultimately determined to be unnecessary additional imaging.

One major factor impacting the performance of breast cancer screening is 2D mammography's greatest limitation: tissue superposition.[7-9] Both screen film and digital

mammography are X-ray projection studies that produce a two-dimensional (2D) radiograph of the three-dimensional (3D) breast. Dense glandular tissue located above and/or below a lesion can occlude or otherwise reduce the visibility of the lesion on the mammogram, and two or more normal features separated in the vertical plane can appear to be a lesion. The former reduces the sensitivity of the study while the latter reduces the specificity. In fact, it has been suggested that tissue superposition alone is responsible for 25% of all mammography recalls.[10]

1.2 Digital Breast Tomosynthesis Imaging

Breast tomosynthesis imaging is a technique that overcomes the tissue superposition issue by acquiring a limited number of X-ray projections from a narrow angular range, and combining these projections to reconstruct a pseudo-3D image.[11] This can provide greater detail about the structure of the tissue and has been shown to reduce the masking effects of tissue superposition.[12-22] The geometry of DBT imaging is very similar to that used in standard mammography, with the difference being that the X-ray tube rotates around the compressed breast, acquiring projections at pre-determined X-ray tube positions (Figure 1.1). During the acquisitions, the detector can either be static or rotate by an amount not necessarily equal to the rotation of the X-ray tube.[23, 24] The set of projections is then used to reconstruct the 3D volume. Several reports on laboratory findings[12-18] and clinical studies[10, 25-30] have suggested that DBT could increase sensitivity and/or specificity as compared to mammography.[10, 25-30] DBT imaging has been in clinical use in Europe and has been FDA approved in the US. It is expected that DBT will be quickly adopted by the radiology community to enhance early detection and diagnosis of breast cancer, with the potential for it to eventually replace mammography as the sole technology used for breast cancer screening of the general population.[10, 31, 32]

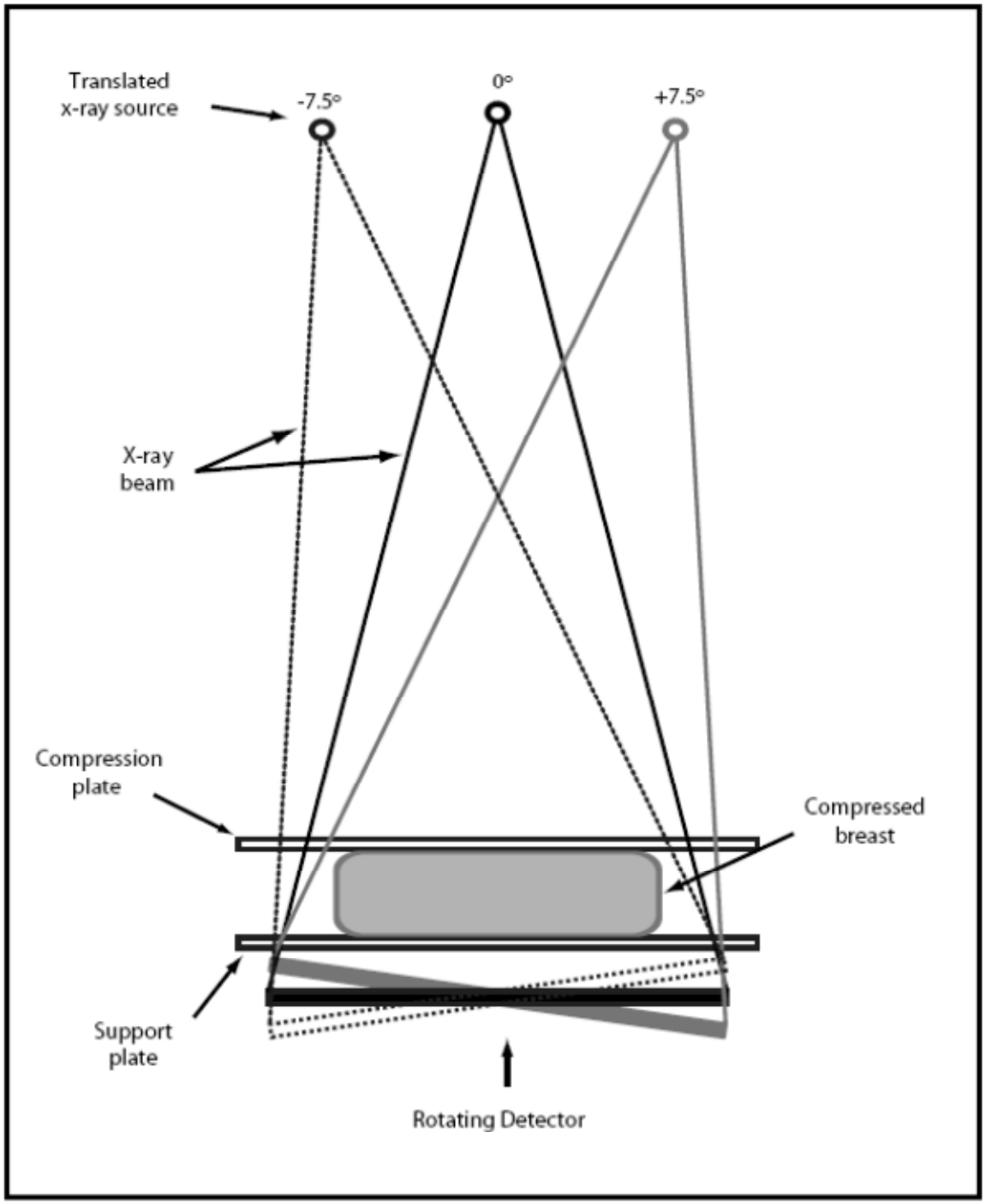

Figure 1.1 – The geometry of digital breast tomosynthesis (DBT) imaging resembles that of mammography, with the difference being that the X-ray tube rotates around the compressed breast and acquires projections over a limited angular range. These projections are then combined into a pseudo-3D image. Some DBT systems utilize a rotating detector such as the one depicted here.

As of this writing, only one DBT system has been approved by the FDA for clinical use in the U.S.: the Selenia Dimensions (Hologic Inc., Bedford, MA). While previous studies have investigated the normalized mean glandular dose for DBT overall, the comprehensive dosimetric characterization study outlined in chapter 3 of this thesis investigates the dosimetry characteristics of this specific tomosynthesis system. It takes into account the specific X-ray spectra and acquisition protocols automatically chosen by the system, and details the absolute mean glandular dose values for breasts of differing thickness and glandular composition. In addition, the Selenia Dimensions is also able to perform full field digital mammography (FFDM), which allows for a direct comparison of dose resulting from FFDM and DBT. The dosimetric characterization of DBT acquisitions by this system also allow for the investigation of novel acquisition techniques, such as that proposed in chapter 6 of this thesis.

However, there are at least five other systems currently in development that could arrive in the clinic over the next several years.[33] Some of these DBT systems incorporate design aspects to differentiate them from the geometry shown in Figure 1.1. For instance, some systems replace the conventional X-ray tube with a carbon nanotube array to avoid focal spot blurring due to X-ray tube motion.[34-40] Such a system could also potentially shorten total acquisition time. Another DBT system incorporates a scanning slit photon counting detector and a collimated X-ray fan beam, resulting in low scatter signal and electronic noise and high quantum efficiency.[41-43] Others have proposed a DBT system in which the X-ray tube is not limited to a single movement plane during acquisition. There are designs in which the X-ray source and detector move in circular paths parallel to each other,[44] and also those which incorporate an arc-and-line CT scan to improve the resolution quality of the DBT acquisition[45]. Research has also been performed on systems in which the X-ray source moves in two arcs normal to each other to cover a portion of a spherical surface above the imaged breast, to improve image quality by sampling the frequency domain to a greater extent. [46, 47]

1.3 Approach - Novel Software-Based Solutions

The emergence of DBT as a potential replacement or adjunct to mammographic screening mandates that solutions be found to two of its major limitations, namely X-ray scatter and mono-energetic reconstruction methods. All the previously described tomosynthesis systems that incorporate a full-field detector lack X-ray scatter reduction measures, be it in software or hardware, leading to the inclusion of the entirety of the X-ray scatter signal in the tomosynthesis projections. In mammography, an anti-scatter grid focused on the fixed X-ray source lowers the ratio of the number of scattered X-rays incident on the detector to that of the number of incident primary (non-scattered) X-rays, or the scatter-to-primary ratio (SPR). In DBT, multiple projections are acquired as the X-ray source rotates around the breast, which prevents the use of such an anti-scatter grid due to the severe cut-off of primary X-rays by the grid in the non-normal projections. It has previously been shown that in DBT the SPR can be as high as 1.6, depending on, in decreasing order of importance, breast thickness, projection angle, breast tissue type, and breast size.[20] It has also been shown that if the scatter X-rays are not removed or otherwise accounted for before tomosynthesis reconstruction, the resulting volume exhibits cupping artifacts and a reduction in the accuracy of reconstruction values of up to 28%.[48-50] Wu et al. also found that inclusion of the scatter signal led to reductions in the signal-difference-to-noise ratio (SDNR) of up to 60%, and in the contrast of lesions of up to 30%.[49]

To reduce the impact of scattered X-rays in tomosynthesis, I investigated the feasibility of a software-based X-ray scatter correction algorithm, centered on Monte Carlo (MC) simulations of a DBT system, which is applied to the acquired tomosynthesis projections before reconstruction. This approach was devised by Dr. Ioannis Sechopoulos, and is based on pre-computed MC simulations of the scatter signal in acquired tomosynthesis projections.[20] Once a map of the scatter signal in a projection of a standard shaped breast is identified, it undergoes deformation and registration to match

the image of the patient breast. This is then used to determine the primary signal of the acquired projection. These estimated primary images are then to be used in the reconstruction process, in place of the acquired projections.

A software-based approach allows for the solution to be implemented without necessitating the addition or modification of any hardware to DBT systems that have already been FDA-approved. This approach also does not modify the image acquisition process itself in any way; therefore there is no increase in resulting radiation dose to the patient, nor is there any change in acquisition time. This would prevent the patient from being subjected to any additional discomfort or pain. I have optimized and applied this algorithm to DBT acquisitions of both phantoms and patients, and have studied the resulting image reconstructions. In chapter 2 of this thesis, I describe this algorithm in further detail and also present a study comparing the image quality of the original uncorrected reconstructions to that of the scatter corrected reconstructions. I have determined that this approach holds great potential to improve the clinical performance of DBT.

This X-ray scatter correction algorithm did require further research to bring it closer to use in the clinical realm. In chapter 7 of this thesis, I describe the development of a model of the scatter images of homogenous clinically-encountered breasts of various sizes to optimize the application of the scatter map in the scatter correction algorithm. This scatter model will be used to generate a set of scatter maps which can then be used in the algorithm to scatter correct any acquired DBT image without the need for a pre-computed library of MC simulations that encompasses the range of breast shapes and sizes that may be encountered clinically.

The reconstruction of the pseudo-3D image from the set of projections of a DBT acquisition is a crucial step which is limited by the false assumption of current reconstruction methods that the energies of X-rays used in the acquisition beam are identical, and thus all undergo the same attenuation as they pass through the breast tissue.

The three main types of reconstruction algorithms all suffer from this same limitation.[15] This mono-energetic assumption is adequate for applications such as computed tomography (CT), which utilizes a heavily filtered, high energy X-ray beam. However, X-ray beams typically used in DBT imaging include a wider and lower range of energies (Figure 1.2), making them more susceptible to beam hardening. The distribution of X-ray energies that reaches the image detector differs greatly from that of the beam entering the breast after undergoing attenuation in the tissue. The only way to account for this phenomenon is to reconstruct the DBT 3D images using a revolutionary spectral reconstruction algorithm, such as the one developed by Dr. Sechopoulos, Dr. James Nagy, and Dr. Julianne Chung that takes into account the distribution of X-ray energies present in the beam and the different attenuation that X-rays of different energies undergo.[51]

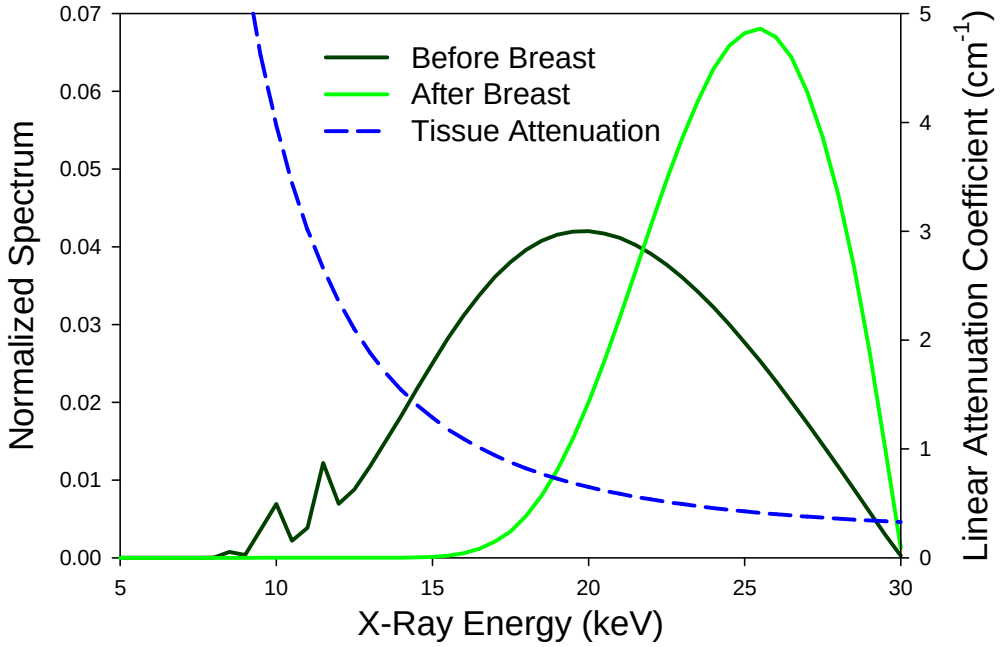

Figure 1.2 – Typical X-ray spectra used in BT imaging, upon entering (black) and exiting (green) the imaged breast. Both spectra are normalized to an area of unity. The variation in spectrum throughout the breast is apparent. The blue dashed line depicts the linear attenuation coefficient of breast tissue, which can be seen to vary considerably with X-ray energy.

This spectral reconstruction algorithm takes as input the distribution of X-ray energies, or X-ray spectrum, present in the beam which allows for a better representation of the true physical processes of a DBT acquisition. The algorithm reconstructs a pseudo-3D image that is a more accurate representation of the breast, and offers the potential to allow for new image acquisition techniques, such as DBT acquisition with varying X-ray beams. This may result in an increase in early breast cancer detection and a reduction in unnecessary additional procedures to investigate false positives.

In chapter 6, I describe the optimization and testing of a new implementation of the spectral reconstruction, collaboration of our group and Dr. James Nagy's group, targeted for graphics processing units (GPUs). This GPU implementation, developed in C++ using OpenCL libraries, greatly shortens the computational time, to clinically feasible limits, and is fully detailed in Bustamante et al.[52] This completed and optimized algorithm implementation was used to reconstruct patient DBT acquisitions, and the image quality of the reconstructions was evaluated and compared to that of mono-energetic reconstructions of the same dataset. In addition, I explored a novel dual-spectrum, single-acquisition imaging technique, possible only with the spectral reconstruction, which has the potential to improve image quality and/or reduce the resulting radiation dose to the patient.

I believe that this dual-faceted approach to improving the image quality of DBT images, firstly by correcting for X-ray scatter and secondly by reconstructing the 3D image using a more physically appropriate method, is very innovative. A scatter correction algorithm centered on an analytical model of X-ray scatter and primary images is novel and distinct from anything that has been previously attempted. In addition, the comprehensive analysis of the range and variation in the size of patient breasts in both the cranio-caudal and medio-lateral oblique views during mammographic and DBT imaging, described in chapter 4, will prove useful in future research in dosimetry, image registration and computer-aided diagnosis. The spectral reconstruction method is novel

and distinct from all current mono-energetic reconstruction methods for DBT imaging. In addition, the dual-spectrum single-acquisition DBT technique has, to my knowledge, not yet been attempted.

This thesis in total documents the successful completion of the development, optimization, and testing of the scatter correction and spectral reconstruction algorithms. These provide improved DBT image quality, are ready for implementation and clinical evaluation in a DBT system, and have the potential to increase the sensitivity and specificity of breast cancer detection and diagnosis with breast tomosynthesis. The feasibility of a new dual-spectrum, single-acquisition, DBT imaging technique has also been evaluated and found to demonstrate great potential in improving image quality while reducing patient dose.

CHAPTER 2

A SOFTWARE-BASED X-RAY SCATTER CORRECTION METHOD FOR BREAST TOMOSYNTHESIS

In this chapter, I present a software-based scatter correction method for digital breast tomosynthesis (DBT) imaging, and an investigation into its impact on the image quality of tomosynthesis reconstructions of both phantoms and patients.

A Monte Carlo (MC) simulation of X-ray scatter was developed and used to generate maps of the scatter-to-primary ratio (SPR) of a number of homogeneous standard-shaped breasts of varying sizes. Dimension-matched SPR maps were then deformed and registered to DBT acquisition projections, allowing for the estimation of the primary X-ray signal. Noise filtering of the estimated projections was then performed to reduce the impact of the quantum noise of the X-ray scatter. This process was tested on acquisitions of a tomosynthesis phantom and on acquisitions of patients. The image quality of the reconstructions of the scatter-corrected and uncorrected projections was analyzed.

The reconstructions of the scatter-corrected projections demonstrated superior image quality. Some embedded microcalcifications in the tomosynthesis phantoms were visible only in the scatter-corrected reconstructions. The visibility of the findings in patient images was also improved by the application of the scatter correction algorithm. The spatial resolution of the images was not adversely affected. This software-based scatter correction algorithm exhibits great potential in improving the image quality of DBT acquisitions of both phantoms and patients.

2.1 Significance

Digital breast tomosynthesis (DBT) imaging[11] is emerging as a potentially feasible replacement or adjunct to standard mammography for the screening and diagnosis of breast cancer. This imaging method allows for the generation of pseudo-three-dimensional (3D) images of the breast, which can provide greater detail about the structure of the tissue, and overcomes one of the main limitations of mammography, that is the compression of information of a 3D structure into a two-dimensional image. In DBT, multiple low dose projections of the breast are acquired over a limited angular range; these projections are then used to calculate a pseudo-3D reconstruction of the imaged tissue.[11, 19, 53] Preliminary studies have shown that DBT has the potential to improve on the clinical performance of breast cancer screening and/or diagnosis.[25-28]

Current tomosynthesis systems, including the Selenia Dimensions (Hologic Inc., Bedford, MA) used in this study, lack X-ray scatter reduction measures, be it in software or hardware, leading to the inclusion of the entirety of the X-ray scatter signal in the tomosynthesis projections. In mammography, an anti-scatter grid oriented towards the fixed X-ray source lowers the ratio of the number of scattered X-rays incident on the detector to that of the number of incident primary (non-scattered) X-rays, or the scatter-to-primary ratio (SPR). In DBT, multiple projections are acquired as the X-ray source rotates around the breast, which prevents the use of an anti-scatter grid due to the severe cut-off of primary X-rays by the grid in the non-normal projections. Therefore, as we have previously shown, in DBT the SPR can be as high as 1.6.[20] It has also been shown that if the scatter X-rays are not removed or otherwise accounted for before tomosynthesis reconstruction, the resulting volume exhibits cupping artifacts, reduced accuracy in reconstruction values, and reduced contrast. [48-50]

Here we present a software-based X-ray scatter correction algorithm that is applied to the acquired tomosynthesis projections before reconstruction. The strength of

this algorithm lies in a pre-computed "library" of SPR maps for homogeneous standard-shaped breasts of various sizes that can be used to scatter-correct any acquired patient DBT image without performing a case-specific and time-intensive Monte Carlo (MC) simulation. We have applied this method to tomosynthesis acquisitions of both phantoms and patient data, and studied the resulting image reconstructions. By comparing the image quality of the original uncorrected reconstructions to that of the scatter corrected reconstructions, we have determined the effectiveness of our method and its ability to potentially improve the clinical performance of DBT when its use is widespread.

2.2 Materials and Methods

The scatter correction algorithm is comprised of four distinct steps: retrieval of a previously computed Monte Carlo (MC) simulation of a SPR map of a geometrically similar DBT acquisition, registration of the SPR map to the acquired breast projections, removal of the low-frequency scatter signal, and noise filtering of the projections. The retrieved SPR images are computed from a MC simulation implemented in C++ based on the Geant4 Monte Carlo simulation toolkit,[54, 55] similar to that used in Sechopoulos et al.[19-21] The simulations generate two images, comprised of the incident scattered and primary X-ray energy, which are combined to form a scatter-to-primary (SPR) map. This SPR map is then deformed and registered to the acquired breast projections. The registered SPR map and the original projection are used to estimate the primary signal, henceforth known as a scatter corrected projection. Lastly, the scatter corrected projections undergo noise filtration and three dimensional (3D) reconstruction.

This procedure was tested on multiple tomosynthesis projection sets acquired by a clinically approved breast digital tomosynthesis system (Selenia Dimensions, Hologic Inc., Bedford, MA) installed at Emory University. The first tests were performed on images of a heterogeneous background Model 020 BR3D tomosynthesis phantom (CIRS Inc., Norfolk, VA), including a "target" plate with representations of embedded spheroid

masses, fibers, and microcalcification clusters. Subsequently, as a preliminary study, the scatter correction algorithm was performed on acquisitions from two distinct patients, one in which a mass is present, and one in which microcalcifications are present. Given the limited number of cases included in this study, only SPR maps with geometry matching these cases were generated from the MC simulation. In the future, a comprehensive library of SPR maps of standard shaped breasts can be assembled, and the corresponding SPR maps would be selected. After correction, all cases were reconstructed using the Maximum Likelihood-Expectation Maximization (MLEM) reconstruction algorithm,[53] although any reconstruction algorithm could be used.

2.2.1 Monte Carlo Simulations

We developed our C++ implementation of MC simulations to match the acquisition geometry of the Selenia Dimensions system for the cranio-caudal (CC) view. The system acquires 15 projections over a 15 degree angular range, with the X-ray tube positioned directly perpendicular to the image detector for the middle (tomosynthesis angle = 0 deg) projection. The simulation includes the X-ray source, the image detector, the breast support plate, and the breast compression paddle. The X-ray tube is modeled as a point source with the center of rotation located at the surface of the detector. Complete details of the system have been previously published in Ren et al.[56] The source to imager distance (SID) for the center projection is set at 70 cm (Figure 2.1). The detector measures 24 x 29 cm, with a pixel pitch of 70 μm. For each test case, the following parameters are determined from the system and the center projection: X-ray tube voltage (automatically selected by the system), compressed breast thickness (Th) (displayed by the system), chest-to-nipple distance (CND), and length along the chest wall (LCW) (both from the center projection). The compressed breast is modeled as a truncated semi-ellipsoid of a homogenous 50%/50% mixture of adipose and glandular tissue positioned at the chest wall edge of the detector with dimensions similar to the case to be corrected

(Figure 2.2). When simulating breast phantoms as opposed to patients' breasts, the breast model is altered such that it has constant thickness at the outer edge, away from the chest wall, to match the straight edges of the phantoms.

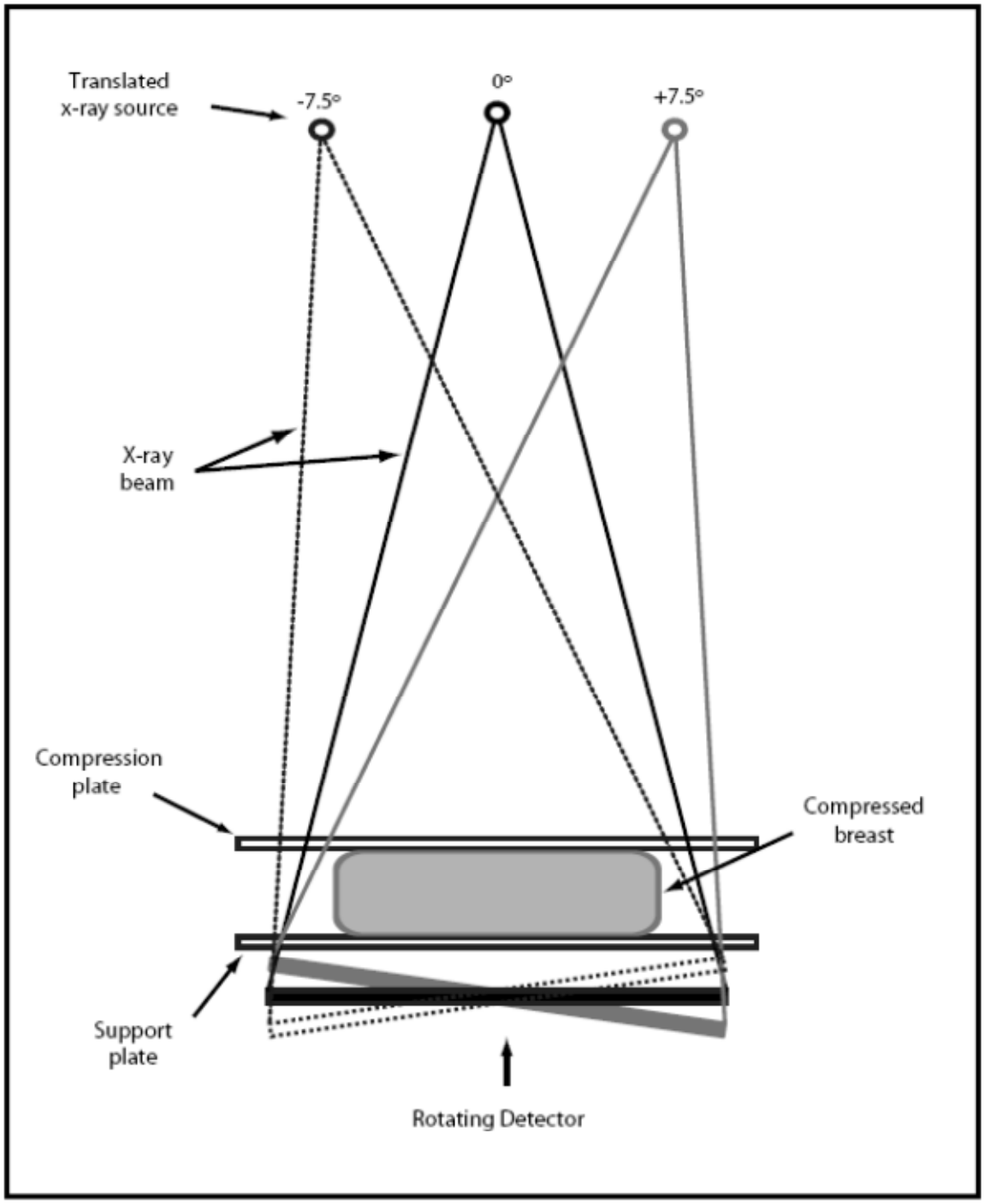

Figure 2.1 – The breast tomosynthesis system acquires 15 projections (of which only 3 are shown here), spaced evenly, from -7.5 degrees to +7.5 degrees of the central line perpendicular to the image detector plane, located at the edge of the detector next to the chest wall

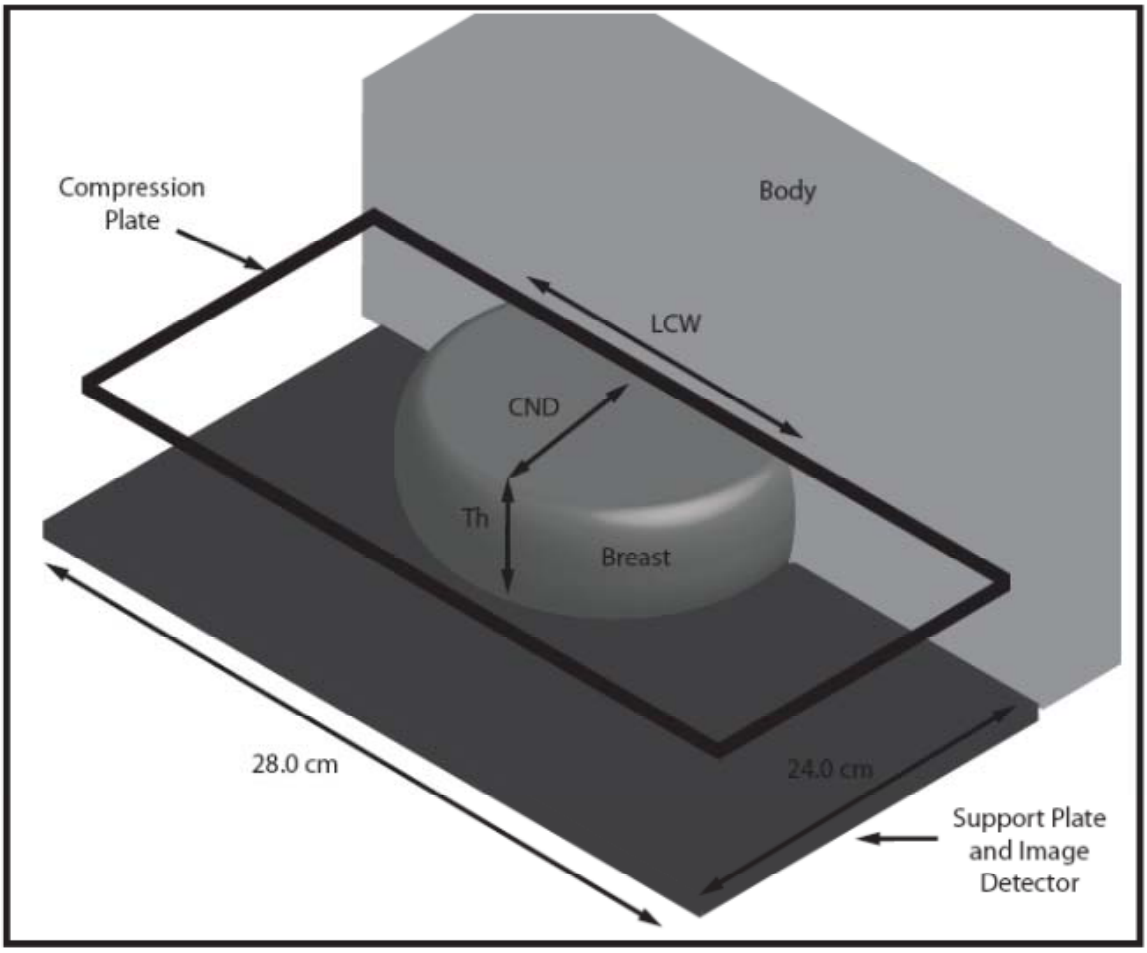

Figure 2.2 – Monte Carlo simulations of the breast in the CC view and system geometry include the image detector, measuring 24 x 28 cm, with 7.0 mm pixels, the support plate, the compressed breast, and the compression paddle. The breast parameters of chest-to-nipple distance (CND), length along the chest wall (LCW), and compressed breast thickness (Th, not depicted) are determined for each case and input as parameters to the simulation. The compressed breast depicted here has curved edges opposite the chest wall, resembling patient breasts. For the simulations for the tomosynthesis phantom, the simulated breast model had straight edges.

For each test case, we performed simulations of five hundred million (5.0×10^8) mono-energetic X-rays emitted from the X-ray tube towards the detector for each energy level from 10.0 keV to the maximum energy level of the spectrum selected by the tomosynthesis system, in 0.5 keV steps. Each photon that arrives at the detector face was recorded as a primary X-ray, or as a scatter X-ray if it experienced at least one Rayleigh

and/or Compton scattering event. Since it is known that X-ray scatter varies slowly over the detector, a pixel pitch of 0.7 mm was used in the MC simulations to decrease the simulation time. The simulated scatter mono-energetic images were combined into estimated scatter spectral images using X-ray spectrum models calculated according to the method described by Boone et al[57] and noise was reduced using the method described by Colijn and Beekman.[58] To obtain noise-less X-ray primary images, raytracing[59] was performed to determine the pathlength of each ray from the source to each pixel on the image detector. The pathlengths of the rays through each material in the simulation, combined with the attenuation coefficients for each material[60] and X-ray energy were used to calculate the estimated primary spectral images. This allows for the calculation of a scatter-to-primary (SPR) map, which depicts the ratio of the scatter signal to the primary signal on a pixel by pixel basis, for each tomosynthesis acquisition projection. As a final step, the SPR map is resampled to the pixel pitch of the imaging system (140 μm).

2.2.2 Deformation and Registration of Scatter-to-Primary Map

In the next step of our scatter correction algorithm, the SPR maps obtained from MC simulations are deformed and registered to the acquired projections to match the breast tissue edge. This is crucial as the MC simulation utilizes a standardized breast shape, so its results have to be matched to the shape of the actual breast images that are being corrected. The small variability of the X-ray scatter signal with glandular fraction and breast size[20] allows for the use of a SPR map of a homogeneous standard-shaped breast of similar, but not equal, size to the one being corrected, after registration to the edges of the acquired image. In this way, a time-consuming MC simulation does not have to be performed for each DBT case acquired.

To perform the registration, the simulated scatter and primary X-ray images are combined to generate a total estimate (T_e) of the acquired projection (T_{acq}). The edge of

the breast and the chest wall in T_{acq} is located by thresholding, using Roberts cross gradient operators, and edge thinning. [61-63] Points along the edge and chest wall are chosen, defining an outer border. For each of these outer points, an inner point is defined 1 cm inwards, parallel to the image gradient, of the breast tissue projection. The outer and inner borders of the breast and the center of mass are used in a thin-plate spline interpolation registration of T_e to T_{acq} in order to determine the transformation (R) necessary to match the shape and position of the of these two breast images.[64] This transformation, consisting of a deformation and rotation and translation, is then applied to the SPR map to generate a registered SPR map (R{SPR}).

For any signal T that consists of the sum of a scatter signal (S) and a primary signal (P), it is known that S, SPR, and T are related according to the equation:

$$S = T\left(\frac{SPR}{1+SPR}\right) \qquad \text{(Equation 2.1)}$$

In this application, as the scatter signal in tomosynthesis has been shown previously to be composed primarily of low frequency signal, T_{acq} needs to be low-pass filtered before it is used in Equation 2.1. Therefore, a low-pass filtered version of T_{acq} ($T_{acq\text{-}LP}$) is used to calculate the registered estimated scatter images (S_e) using the following equation:

$$S_e = T_{acq-LP}\left(\frac{R\{SPR\}}{1+R\{SPR\}}\right) \qquad \text{(Equation 2.2)}$$

The final registered primary X-ray estimates (P_e) are then calculated by subtracting S_e from T_{acq}.

2.2.3 Noise Filtration and Image Reconstruction

Before reconstructing the breast tomosynthesis image, the P_e images undergo noise-filtering to reduce the impact of the quantum noise of the scattered X-rays, since the previous steps only correct for the low-frequency offset from the scatter signal. For this, a 3x3 pixel kernel adaptive means filter is utilized, with a minimum variance level

0.25 times the noise variance level measured from tomosynthesis image acquisitions of matched thickness phantoms representative of homogenous 50%/50% glandular/adipose tissue (CIRS Model 082).[65]

Finally, the DBT reconstruction is performed using the noise filtered P_e. For this study, the maximum likelihood expectation maximization (MLEM) iterative method described by Wu et al was used, as this method has been shown to be very effective at balancing the image quality of both masses and microcalcifications.[15, 53] However, the corrected projections resulting from our method can be used for reconstruction with any algorithm.

2.2.4 Image Quality Comparison

In order to determine the benefits realized by the scatter correction algorithm, the complete process described above was performed on acquisitions of a phantom and patient cases, and the image quality of the resulting scatter-corrected reconstructed 3D images was compared to that of the original, uncorrected reconstructions using image contrast metrics such as maximum signal difference, signal-difference to-noise ratio, and integrated mass signal. In addition, to better understand the impact that each portion of the scatter correction algorithm has on the reconstructed image quality, reconstructions of the phantom projections after undergoing only the X-ray scatter quantum noise filtration (T_F) were also obtained and evaluated. The datasets are all reconstructed to image stacks of 0.14 mm x 0.14 mm x 1.0 mm voxels.

The patient case datasets were acquired for an unrelated IRB-approved clinical study from which the images were released for use in other research projects. For the phantom cases, the CIRS Model 020 BR3D Tomosynthesis phantom was utilized, consisting of 1 cm thick slabs of a heterogeneous 50%/50% mixture of adipose and glandular tissue. This phantom represents real tissue better than a homogenous phantom, and tests tomosynthesis' ability to suppress tissue superposition. One of the 1 cm thick

slabs, the “target slab”, has embedded spheroid tumor masses, fibers, and microcalcification clusters of differing sizes (Figure 2.3). Phantoms measuring 5 cm and 8 cm thick were imaged, with the target plate located at heights of 3 cm and 5 cm, respectively, and the image quality of the 3D images was evaluated using three metrics.

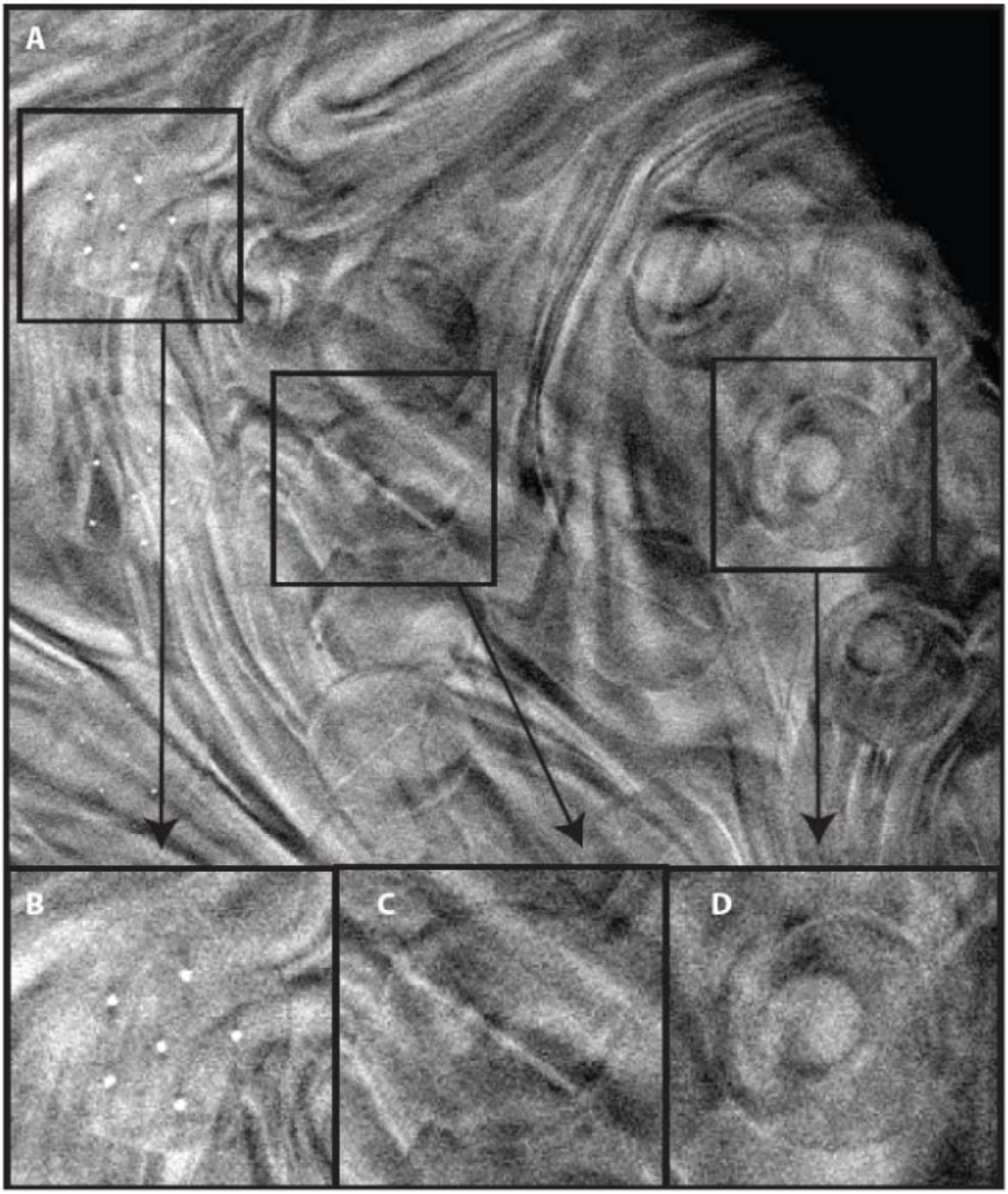

Figure 2.3 – (a) The CIRS Model 082 Breast Mammography Phantom, representing a 50/50 heterogeneous mixture of glandular and adipose tissue contains three types of embedded targets of varying sizes: (b) spheroid masses, (c) fibers, and (d) microcalcification clusters.

The maximum signal-difference (SD) and signal-difference to noise ratio (SDNR) of the spheroid mass targets were calculated by measuring the maximum and minimum signal of a region of interest (ROI) centered on each mass and the noise of a designated area of background signal in the focal plane. The circular ROI around the spheroid mass was chosen such that its diameter spanned 60 voxels and encompassed each mass entirely. Therefore, the SD and SDNR were calculated as follows:

$$SD = I_{max} - I_{min} \tag{3}$$

$$SDNR = \frac{I_{max} - I_{min}}{\sigma_{BG}}, \tag{4}$$

where I_{max} is the maximum voxel value in the ROI, I_{min} is the minimum voxel value in the ROI, and σ_{BG} is the standard deviation of voxel values in the background area, which is a fixed circular region with a diameter of 25 voxels in the focal plane, chosen for its homogeneity.

Another quantitative measure of the image quality of the spheroid mass targets is the gray level signal profile measured in the direction perpendicular to the chest wall. From this profile, the integrated mass signal (IMS) was calculated for each mass. We define IMS as the area under the signal profile curve above the signal at the edge of the mass itself, where the minimum gray levels are found (Figure 2.4). The IMS therefore represents the combined signal difference of the mass itself from its immediate surroundings.

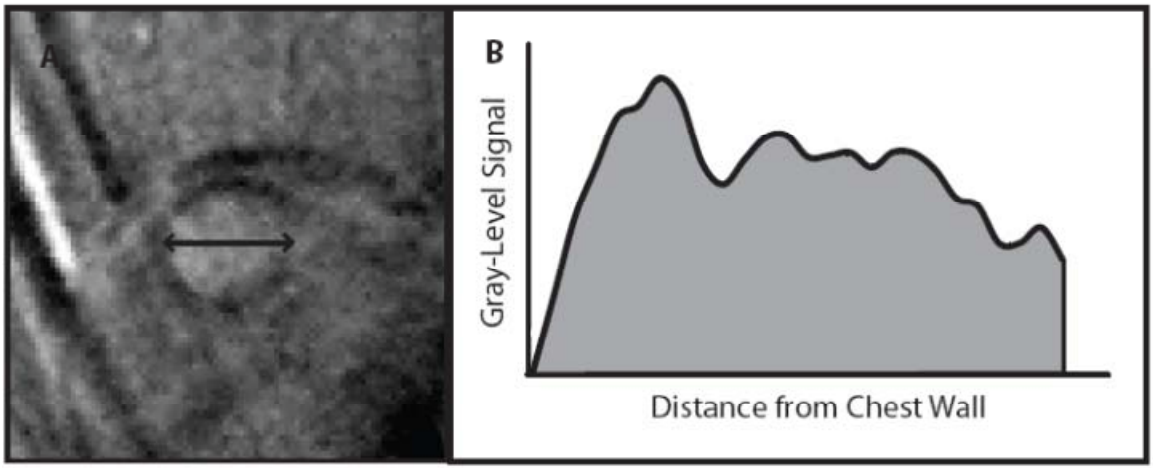

Figure 2.4 – (a) To measure the Integrated Mass Signal (IMS), between the mass edges, gray level signal profiles of each of the spheroid masses were taken in the direction perpendicular to the chest wall. (b) The gray level signal is offset from the minimum signal level of the profile, which is found at the edge of the spheroid mass. The IMS is the area (grey) under the offset profile curve.

To determine if the scatter correction algorithm has a significant impact on the image contrast of the spheroid masses, the Wilcoxon signed-rank test was used. The test was applied to the SD, SDNR, and IMS for all masses in both the 5 cm and 8 cm thick phantoms, to compare differences in each measure between the original reconstructions and either test case: (i) the reconstructions of the scatter corrected projections or (ii) the reconstructions of noise filtered-only projections. In both cases, the null hypothesis is that the median difference of the image contrast measurement, across the masses, between the original reconstruction and the test case is zero. The Wilcoxon signed-rank test was chosen over a paired t-test as the distribution of differences could not be assumed to be normally distributed.

To ensure that the resolution of the DBT reconstructions was not compromised by the noise filtration process, the modulation transfer function (MTF) of the projections before and after undergoing filtering was measured. The line response function (LRF) of the projections of the system was measured using a TX-5 tungsten edge (Scanditronix Wellhöfer, Schwarzenbruck, Germany), following the method of Kyprianou et al.[66] The

MTF was calculated by taking the Fourier transform of the LRF.[67] These projections then underwent noise filtration with the filter's variance parameter set to 1, 0.5, and 0.25 times the noise variance level, the lowest level (0.25x) being the actual level used in the noise filtration process. The MTF of the noise filtered projections was compared to that of the unfiltered projections.

2.3 Results

2.3.1. Scatter Correction of Phantom Images

Reconstructed slices of the focal plane of the target plate of the 5 cm and 8 cm thick phantoms are shown in Figure 2.5 and Figure 2.6, respectively. Figure 2.5(a) depicts the reconstruction of the original uncorrected projections (*O*), while Figure 2.5(b) shows a zoomed-in view of the second largest of the spheroid masses (M2) and Figure 2.5(c) shows a zoomed-in view of the 4th largest microcalcification cluster (C4). Figure 2.5(d) - Figure 2.5(f) depict the same ROIs, respectively, of the reconstruction of the noise filtered T_F projection set (*F*) , while Figure 2.5(g) - Figure 2.5(i) depict the same ROIs, respectively, of the reconstruction of the fully scatter corrected projection set (C) of the 5 cm thick phantom. All images are displayed with equal window width but different window levels. In Figure 2.5(i), white arrows point to two microcalcifications of the C4 set which can be clearly seen in the *C* reconstruction, but not in the *O* reconstruction (Figure 2.5(c)) and to a lesser extent in the *F* reconstruction (Figure 2.5(i)). Figure 2.6 depicts similar views of the three reconstructions of the 8 cm thick phantom, with the zoomed in views centered on the fourth largest spheroid mass (M4) and C4. In Figure 2.6(i), white arrows point to two microcalcifications of the C4 set which, like in the 5 cm phantom, can be clearly seen in the *C* reconstruction, but not in the *O* reconstruction (Figure 2.6(c)) and to a lesser extent in the *F* reconstruction (Figure 2.6(f)).

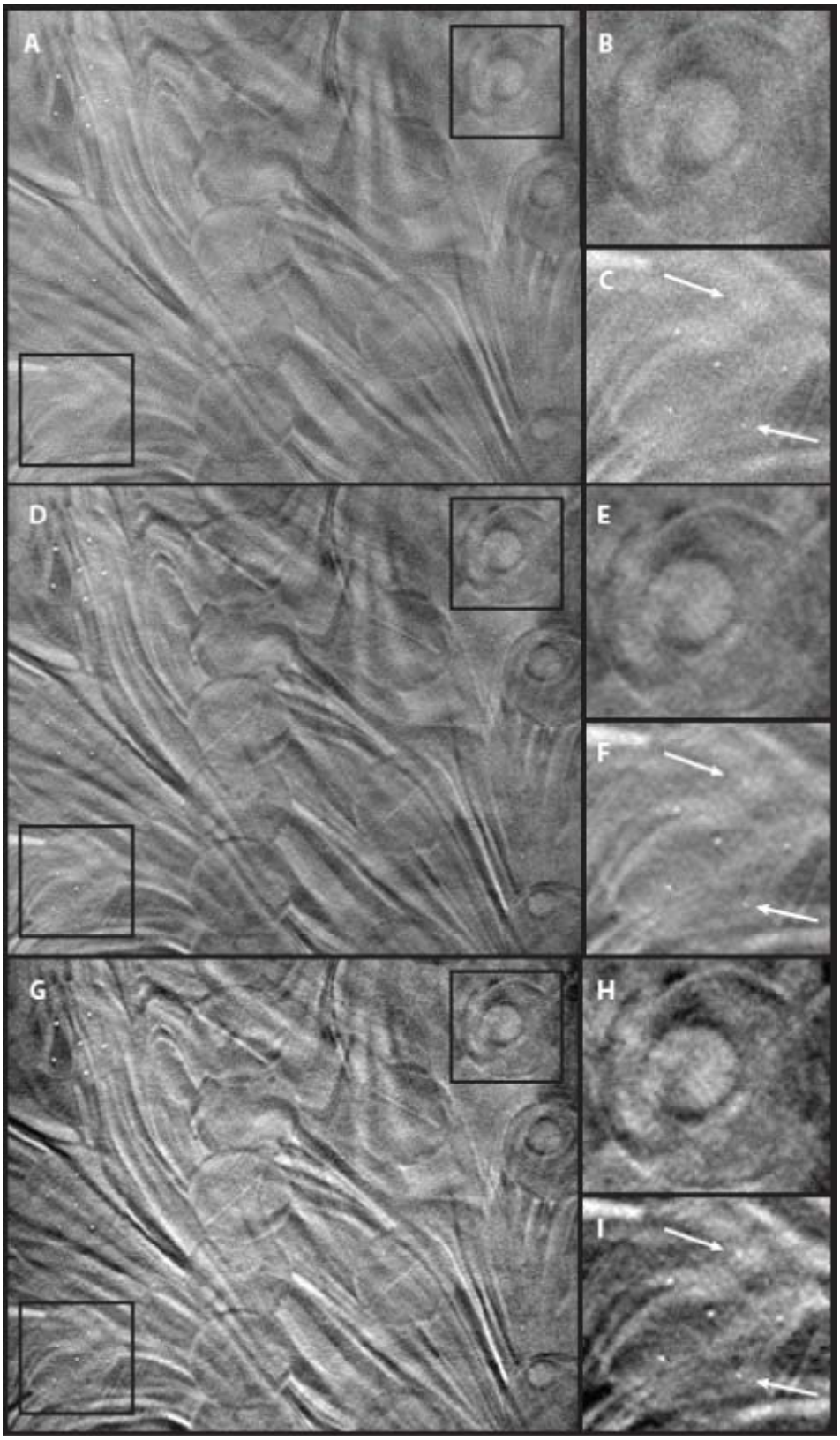

Figure 2.5 – (a) Uncorrected in-plane slice of the focal plane of the embedded target plate located at the center of the 5 cm thick heterogeneous phantom, with marked regions

of interest (ROI). (b) Uncorrected zoomed-in view of the ROI centered on the second largest of the spheroid masses (M2). (c) Uncorrected zoomed-in view of the ROI centered on the fourth largest microcalcification cluster (C4). White arrows point to microcalcifications that are present, but not visible. (d) Scatter corrected slice of the embedded target plate of the 5cm thick phantom, with boxed ROIs. (e) Scatter corrected zoomed-in view of the ROI centered on M2. The SDNR of this mass exhibited a 65% increase following scatter correction. (f) Scatter corrected zoomed-in view of the ROI centered on C4. White arrows point to microcalcifications not visible in (c) that are visible here. These images are displayed with equal window width but different window levels.

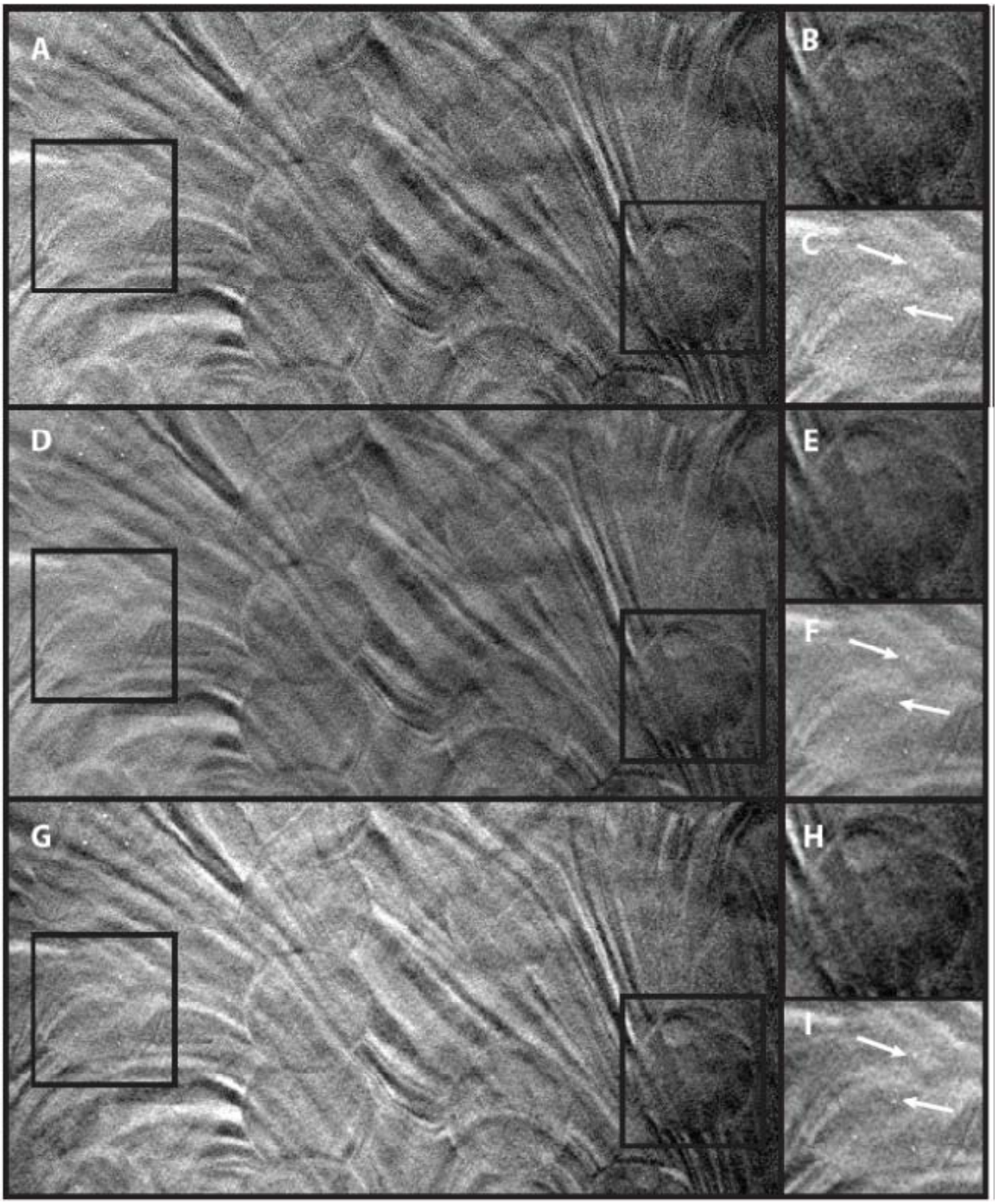

Figure 2.6 – (a) Uncorrected in-plane slice of target plate at center of 8 cm thick phantom, with boxed ROIs. (b) Uncorrected zoomed-in view of the ROI centered on the fourth largest spheroid mass (M4). (c) Uncorrected zoomed-in view of the ROI centered on C4. White arrows point to microcalcifications that are present, but not visible. (d) Scatter corrected slice of target plate of the 8cm thick phantom, with boxed ROIs. (e) Scatter corrected zoomed-in view of the ROI centered on M4. The SDNR of this mass exhibited a 26% increase following scatter correction. (f) Scatter corrected zoomed-in view of the ROI centered on C4. White arrows point to microcalcifications not visible in

(c) that are visible here. These images are displayed with equal window width but different window levels.

Quantitative analysis of the reconstructions of the *O, F,* and *C* sets revealed that the scatter correction algorithm improves the SD (Table 2.1), the SDNR (Table 2.2), and the IMS (Table 2.3) of the spheroid masses located in the target plate. The *C* reconstructions of the 5 cm thick phantom exhibited a 60%-66% improvement in SD and SDNR across the four largest spheroid masses (M1-M4). The *C* reconstructions of the 8 cm thick phantom exhibited a more modest improvement (17%-32%) in SD and SDNR. The improvements in both metrics were statistically significant (SD: Wilcoxon's $W=0$, $p < 0.01$, $n=8$; SDNR: $W=0$, $p < 0.01$). Note that the *C*/*O* ratio of SD is very similar to the *C*/*O* ratio of SDNR because the measured noise of the *C* reconstructions approaches that of the *O* reconstructions. Although in most cases noise filtering only improves image quality, the fully scatter-corrected reconstructions exhibit the greatest improvement in both SD and SDNR. The *C* reconstructions also demonstrate improvements of 15-62% in IMS ($W=0$, $p < 0.01$), whereas the *F* reconstructions exhibit decreased IMS measurements ($W=0$, $p < 0.01$). This can also be seen in the offset signal profiles displayed Figure 2.7, from which the IMS is calculated. Note also that the *C/O* ratios of IMS measurements for all four spheroid masses of the 8 cm thick phantoms are greater than the corresponding *C/O* ratio of IMS for the 5 cm thick phantom. The two smallest masses (M5 and M6) cannot be clearly seen in any of the reconstructions, and therefore were excluded from the analysis.

Table 2.1 – Ratio of Maximum Signal Difference (SD) of Spheroid Masses in Heterogeneous Tomosynthesis Phantom between the original reconstructions (O) and the noise-filtered-only (F) and scatter-corrected (C) reconstructions. Wilcoxon signed-rank test applied across masses M1-M4 in both phantoms revealed that the differences seen in SD between the C reconstructions and the O reconstructions was significant (Wilcoxon's $W=0$, $p < 0.01$, $n=8$), while the differences in SD between the F reconstructions and the O reconstructions was not ($W=14$, $p > 0.2$).

Table 2.1

Phantom Thickness (cm)	Reconstruction Set	M1	M2	M3	M4
5	*F/O* Ratio	1.10	1.15	1.09	1.19
5	*C/O* Ratio	1.60	1.65	1.66	1.65
8	*F/O* Ratio	0.75	0.83	0.82	0.78
8	*C/O* Ratio	1.17	1.31	1.32	1.26

Table 2.2 – Ratio of Signal-Difference to Noise Ratio (SDNR) of Spheroid Masses in Heterogeneous Tomosynthesis Phantom between the original reconstructions (O) and the noise-filtered-only (F) and scatter-corrected (C) reconstructions. The differences seen in SDNR between the C reconstructions and the O reconstructions and between the F and O reconstructions were both significant (Wilcoxon's $W=0$, $p < 0.01$).

Table 2.2

Phantom Thickness (cm)	Reconstruction Set	M1	M2	M3	M4
5	*F/O* Ratio	1.46	1.53	1.46	1.59
5	*C/O* Ratio	1.60	1.65	1.66	1.65
8	*F/O* Ratio	1.13	1.24	1.23	1.17
8	*C/O* Ratio	1.17	1.31	1.32	1.26

Table 2.3 – Ratio of Integrated Mass Signal (IMS) of Spheroid Masses in Heterogeneous Tomosynthesis Phantom between the original reconstructions (O) and the noise-filtered-only (F) and scatter-corrected (C) reconstructions. The differences seen in IMS between the C reconstructions and the O reconstructions and between the F and O reconstructions were both significant (Wilcoxon's $W=0$, $p < 0.01$).

Table 2.3

Phantom Thickness (cm)	Reconstruction Set	M1	M2	M3	M4
5	*F/O* Ratio	0.85	0.75	0.88	0.91
5	*C/O* Ratio	1.19	1.15	1.22	1.29
8	*F/O* Ratio	0.79	0.90	0.93	0.83
8	*C/O* Ratio	1.26	1.28	1.62	1.30

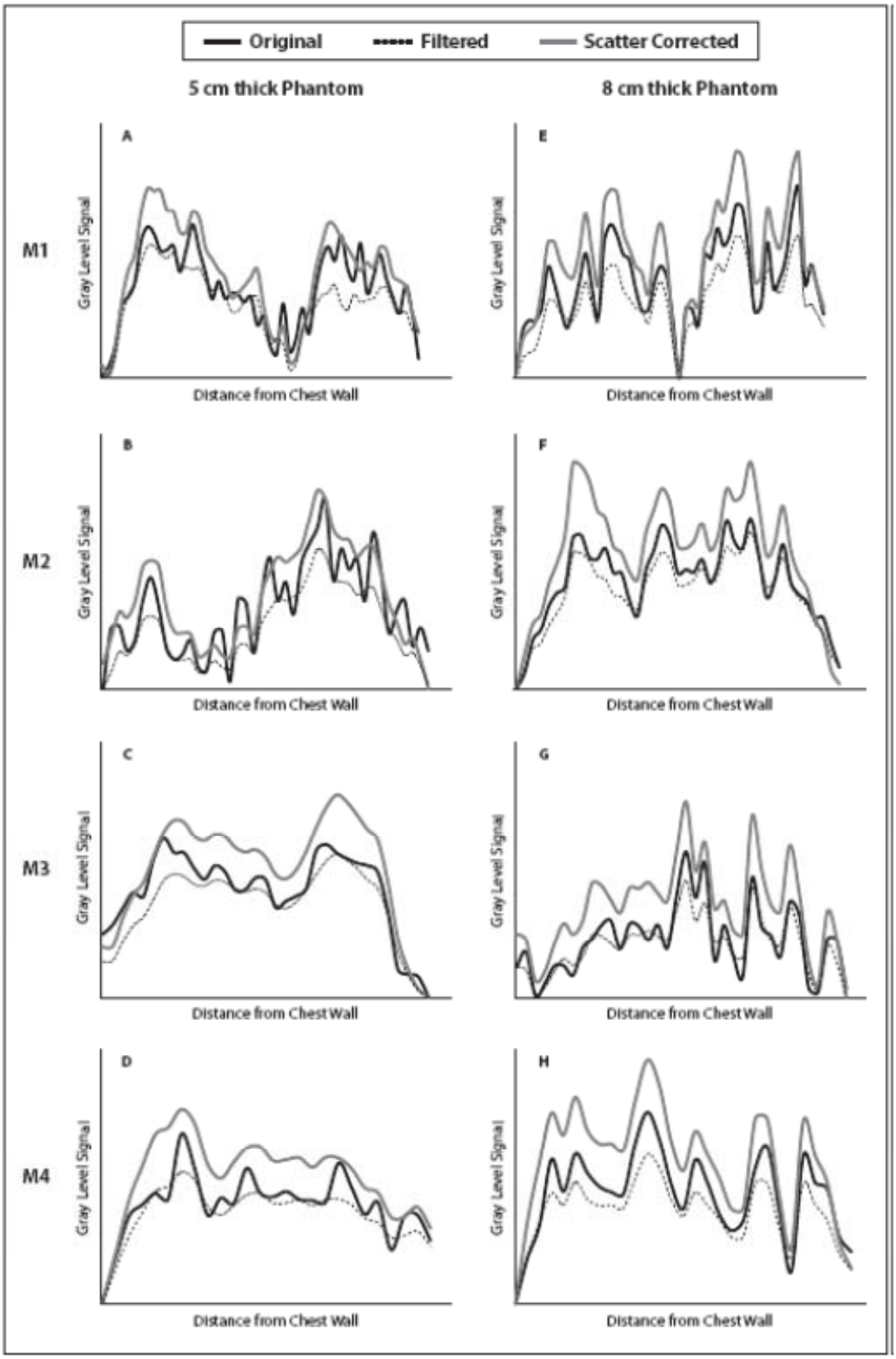

Figure 2.7 – Offset Gray Level Signal Profiles of Spheroid Masses measured from Original (O), Filtered (F), and Scatter Corrected (C) reconstructions – The signal profiles,

measured in the direction perpendicular to the chest wall, of the masses M1-M4 of the target plate of the 5 cm thick phantom can be seen in (a), (b), (c), and (d), respectively, while those of the 8 cm thick phantom can be seen in (e), (f), (g), and (h). The area under the profile curve, or Integrated Mass Signal (IMS), of the C reconstructions (grey) is greater than that of the O reconstructions (black), which in turn is greater than that of the F reconstructions (dashed black). This suggests that the masses are easiest to distinguish from their surroundings in the C reconstructions.

The MTF of the central projection of a DBT acquisition was computed and can be seen in Figure 2.8. The MTF of the system's projections did not differ before and after four degrees of noise filtration was applied to the projections (Figure 2.8). Thus, the noise filtration did not alter the resolution of the Selenia Dimensions system.

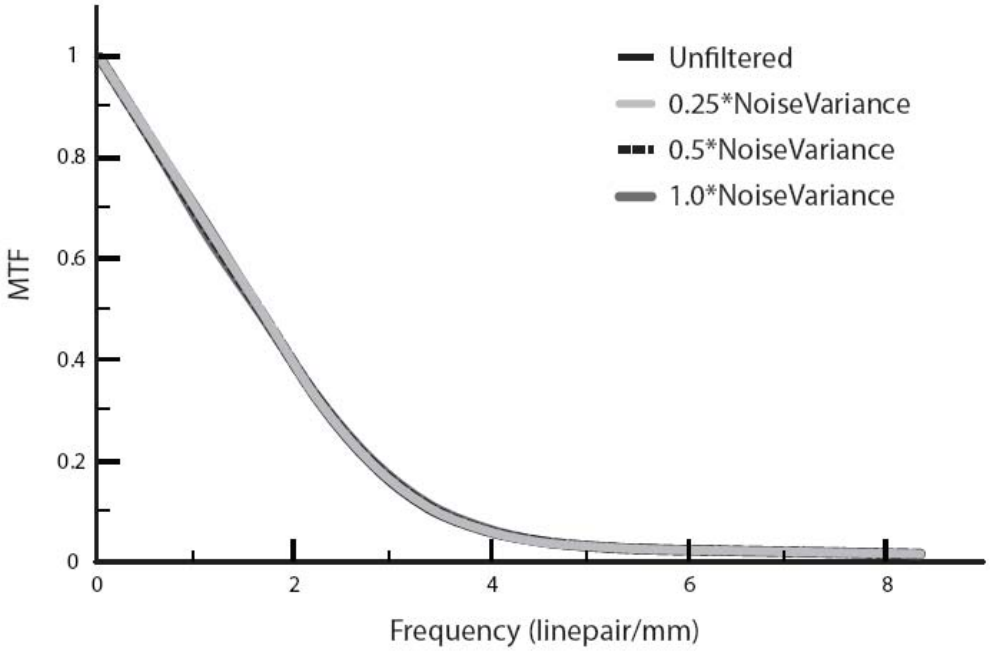

Figure 2.8 - Comparison of MTF before and after noise filtration - The MTF of the projections in the direction parallel to the X-ray tube movement before and after undergoing differing levels of noise filtration. The filter's variance parameter was set to 0.25, 0.5, and 1.0 times the noise variance level.

2.3.2 Scatter Correction of Patient Cases

In-plane views of the *O* and *C* reconstructions of the first patient case (Patient A), centered around a mass anterior to the nipple, are displayed in Figure 2.9(a) and Figure 2.9(b), respectively. In-plane views of the *O* and *C* reconstructions of the second patient case (Patient B), centered on microcalcifications at two different depths, are displayed in Figure 2.10. Figure 2.10(a) and Figure 2.10(b) depict the *O* and *C* sets, respectively, of three microcalcifications located near the X-ray tube side of the breast while Figure 2.10(c) and Figure 2.10(d) depict the same of a single larger microcalcification located near the detector side of the breast. These images are also displayed with equal window width but different window levels.

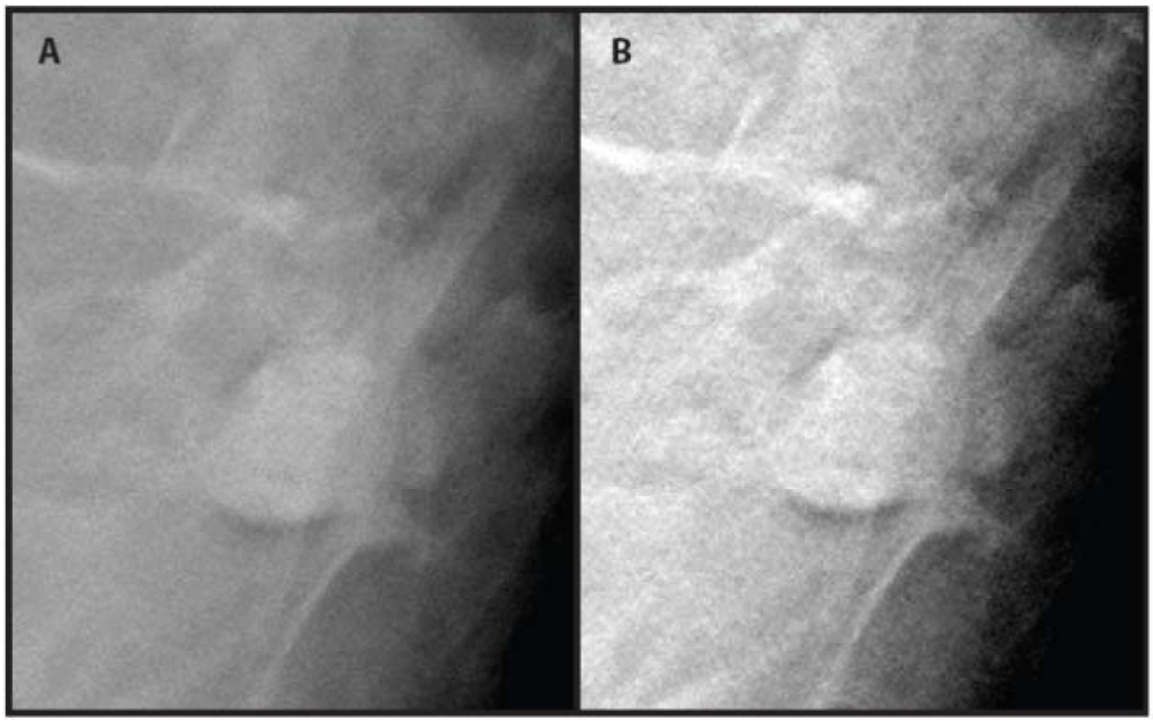

Figure 2.9 – (a) ROI of uncorrected central slice of a mass in Patient A image. (b) Identical ROI of the same slice, following scatter correction of the projections. These images are displayed with equal window width but different window levels.

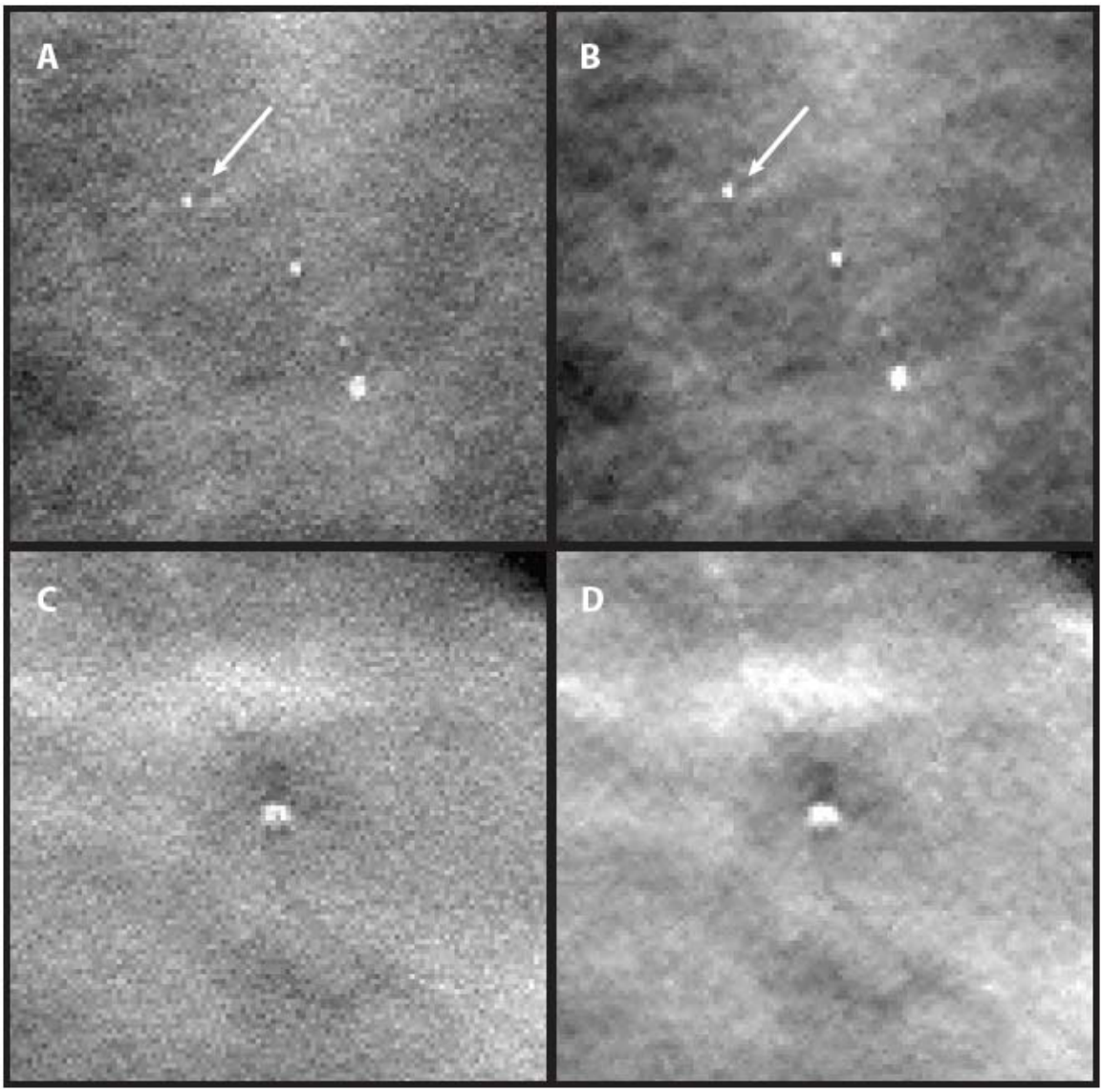

Figure 2.10 - (a) ROI of uncorrected central slice of a microcalcification cluster in Patient B image. White arrow points to the smallest lesion seen in the entire image stack. (b) Identical ROI of the same slice of (a), following scatter correction of the projections. White arrow points to the smallest lesion, more visible after the correction. (c) ROI of uncorrected central slice of a large microcalcification of Patient B in a focal plane separate from (a). (d) Identical ROI of the same slice of (c), following scatter correction of the projections. These images are displayed with equal window width but different window levels.

2.4 Discussion

In this study, we presented a Monte Carlo simulation-based scatter correction algorithm for breast tomosynthesis and compared the image quality of both phantom and patient reconstructions before and after the acquired projections were scatter corrected. Contrast enhancement can be seen in the image reconstructions of the phantoms of differing thicknesses shown in Figure 2.5 and Figure 2.6, where the swirl patterns of the heterogeneous mixture of 50/50 glandular/adipose tissue are more distinguished in the scatter corrected images. The zoomed-in views of the embedded spheroid masses in Figure 2.5 and Figure 2.6, and the improvements in SDNR shown in Table 2.2 and the improvements in IMS shown in Table 2.3,

Phantom Thickness (cm)	Reconstruction Set	M1	M2	M3	M4
5	*F/O* Ratio	1.10	1.15	1.09	1.19
5	*C/O* Ratio	1.60	1.65	1.66	1.65
8	*F/O* Ratio	0.75	0.83	0.82	0.78
8	*C/O* Ratio	1.17	1.31	1.32	1.26

Table both statistically significant, also support this conclusion. While the scatter corrected reconstructions of the 8 cm thick phantom do not exhibit greater improvements in SDNR than those of the 5 cm thick phantom, they do exhibit greater improvements in IMS. As DBT is not a quantitative imaging technique, we do not believe that the SDNR values of the reconstructed images can be reliably compared between different-sized images acquired with differing X-ray spectra. Instead we propose that the IMS, calculated from the offset gray level signal profiles seen in Figure 2.7, allows for a better comparison of the visibility of the spheroid masses between different reconstructions of

the same set of projections. In addition, the zoomed in views of the C4 microcalcification group in Figure 2.5 and Figure 2.6 (white arrows) show that at least two of the microcalcifications can be clearly seen in the scatter corrected reconstructions but not in the original uncorrected set. The fibers also exhibit improved contrast after scatter correction, as they are easier to detect in both phantoms. Improvement in the visibility and detail of the lesions in the patient images were also seen, with no introduction of artifacts.

To determine if the noise filtering used to reduce the impact of the scatter X-rays quantum noise negatively impacts image resolution, we computed and compared the MTF of the central projection. The MTF was not reduced after the central projection underwent various degrees of noise filtration (Figure 2.7). Thus, we believe that there is no change in the spatial resolution of the images following the application of our scatter correction algorithm.

Examination of the patient images also reveals improvements in image quality. The mass in the Patient A image is brighter and more easily identifiable in the scatter corrected reconstructions and greater detail is also visible inside the mass itself (Figure 2.9). After scatter correction, the microcalcifications in the Patient B images shown in Figure 2.10 are also brighter and easier to visualize. In particular, the topmost lesion in Figure 2.10(b) underwent the greatest improvement, which is very encouraging as it is also the smallest lesion.

The improvements in image quality realized by the scatter correction algorithm can not only aid clinicians in breast cancer diagnosis and screening, but it could open avenues for new breast tomosynthesis acquisition techniques. Recent studies have suggested that reduced compression of breast tissue to alleviate pain during tomosynthesis imaging might be possible without compromising image quality and without increasing patient dose.[68, 69] We have performed preliminary studies to evaluate the feasibility of reduced breast compression in breast tomosynthesis with adjustment of

acquisition parameters to maintain or lower tissue dose and have found that image quality might be affected by the increased X-ray scatter signal due to the increased tissue thickness and the change in X-ray spectrum. Subsequent application of the scatter correction algorithm showed great promise in restoring or even enhancing the image quality of the reduced compression acquisitions.

Lastly, this method has been tested extensively on CC views of multiple phantoms and of multiple patient cases acquired by the clinical prototype installed at Emory University. We anticipate no problems in adapting the Monte Carlo simulation geometry to match other tomosynthesis systems or a finalized clinical system. However, we would like to consider improving the representation of the simulated breast tissue. Here, we have used the cut out center of a semi-ellipsoid, generating a symmetrically rounded edge of the breast tissue. In reality, compression of breast tissue probably does not confer such a perfect symmetry to the tissue edge, and could affect the scatter estimates in that region. In addition, work is currently under way to expand the method to accommodate the medio-lateral oblique (MLO) view of DBT systems. While the MC simulation can generate SPR maps of projections acquired in the MLO view, enhancement of the deformation and registration process is required in order to account for the pectoralis muscle included in these projections. We have also shown the potential benefits of this scatter correction method on two distinct clinical acquisitions to demonstrate the possible improvements in image quality on a mass and on microcalcifications. Also, while this is a small sample size, this study does suggest that further investigation of the scatter correction algorithm's ability to improve the image quality of clinical images is warranted. In chapter 5, the effectiveness of this method on clinical DBT acquisitions that contain glandular masses and microcalcifications is explored to a greater degree. The impact of the algorithm on clinical detection will also need to be evaluated in the future.

To bring this method to the clinical realm, the development of an extensive library of scatter maps of breasts of varying sizes and shapes and its use in the algorithm detailed here is also described in chapter 5. In addition, a computational model of scatter maps that can be used to quickly produce a patient-matched map is detailed in chapter 7. The presented algorithm, not including MC simulation time, can be applied in a matter of seconds, which would allow for a seamless inclusion of this technique in clinical DBT. In addition, further work on the noise filtering would include extensively testing multiple variance levels and kernel sizes for use in the adaptive means filter. Our parameters for the noise filtration were chosen once a noticeable improvement in noise level, with minimal impact on sharpness, was observed. Other noise filters might also offer further improvement.

2.5 Conclusion

The software-based scatter correction algorithm presented here was successfully implemented and tested on digital breast tomosynthesis acquisitions of both phantoms and patients. We found measurable improvements in the image quality of the scatter corrected reconstructions of phantoms of two sizes. The application of this method to the reconstructions of the patient images also improved the image quality, including that of actionable findings in a clinical setting. This study has demonstrated the feasibility of a software approach to scatter reduction in DBT and could lead to both improved clinical performance of DBT and faster widespread adoption of this technology.

The promising results of this study also solidify the scatter correction algorithm as the first step towards the goal of this thesis, which is to improve the image quality of DBT imaging.

CHAPTER 3

CLINICAL DIGITAL BREAST TOMOSYNTHESIS SYSTEM: DOSIMETRIC CHARACTERIZATION

This chapter presents a comprehensive characterization of the dosimetric properties of a clinical DBT system for the acquisition of mammographic and tomosynthesis images.

Compressible water/oil mixture phantoms were created and imaged using a Hologic Selenia Dimensions system in both DBT and full field digital mammography (FFDM) mode. Empirical measurements of the X-ray tube output were performed with a dosimeter to measure the air kerma for the range of tube current-exposure time product settings and to develop models of the automatically selected X-ray spectra. A Monte Carlo simulation of the system was developed and used in conjunction with the AEC-chosen settings and spectra models to compute and compare the mean glandular dose resulting from both imaging modalities for breasts of varying size and glandular compositions.

The resulting glandular dose of DBT and FFDM acquisitions is detailed further in this chapter but it is noteworthy that for the 2D/3D fusion imaging performed by the Selenia Dimensions system, the resulting mean glandular dose for a 5 cm thick, 50% glandular breast is 2.50 mGy, which is below the MQSA limit for a two view screening mammography study.

3.1 Significance

The Selenia Dimensions (Hologic Inc., Bedford, MA) DBT system, approved by the FDA for clinical use, is able to perform both full field digital mammography (FFDM) and DBT. Previous studies have investigated the normalized mean glandular dose,

defined as the mean glandular dose per unit reference exposure or air kerma, resulting from DBT overall as an imaging modality.[19, 21, 70, 71] These studies provide relative dose coefficients with units of mean glandular dose per unit reference exposure or air kerma for a number of X-ray spectra. In this study, we investigate the dosimetry characteristics of a specific tomosynthesis system, taking into account the specific X-ray spectra and acquisition protocols automatically chosen by the system to obtain absolute mean glandular dose values for each breast of specific thickness and composition. In addition, the dual imaging functionality of the system allows for a direct comparison of dose resulting from FFDM and DBT. Given the potential for use of this system and the DBT modality for screening of the general population, we believe that in-depth knowledge of its dosimetric behavior is essential.

The purpose of our study was to comprehensively characterize the dosimetric properties of a clinical digital breast tomosynthesis system for the acquisition of mammographic and tomosynthesis images.

3.2 Materials and Methods

To characterize the radiation dose delivered to an imaged breast, the automatic exposure control (AEC) of the DBT system was used to select the X-ray spectrum and X-ray exposure settings for both FFDM and DBT acquisitions of breast phantoms of varying sizes and adipose/glandular compositions. Empirical measurements using a dosimeter were performed to determine the half value layers of the selected X-ray spectra and to measure the air kerma at the breast support plate for different tube current-exposure time product settings. A Monte Carlo simulation of the Selenia Dimensions system was developed and used to compute the normalized mean glandular dose resulting from both mammography (D_gN_{FFDM}) and tomosynthesis (D_gN_{DBT}) acquisitions of breasts of sizes and compositions matching those of our phantoms. Finally, the mean glandular

dose was estimated by combining these results with the tube current-exposure time product settings chosen by the AEC.

3.2.1 Digital Breast Tomosynthesis System

When performing a DBT acquisition, the system acquires 15 projections over a 15 degree angular range (from -7.5 to +7.5 degrees). For the central projection (tomosynthesis angle = 0 deg), the source-to-imager distance (SID) is 70 cm, with a 2.5 cm air gap between the detector and the breast support plate (Figure 3.1). The isocenter of the X-ray tube rotation is located at the central ray of the central projection, on the surface of the detector. The image detector measures 24 cm x 29 cm and rotates about an axis located on the surface of the detector and orthogonal to the chest wall. The system utilizes a tungsten target, with additional aluminum filtration for performing DBT acquisitions and either rhodium or silver filtration for performing FFDM acquisitions. Compression of the breast is achieved with a 3 mm thick compression paddle. Details of this system were previously published by Ren et al.[72]

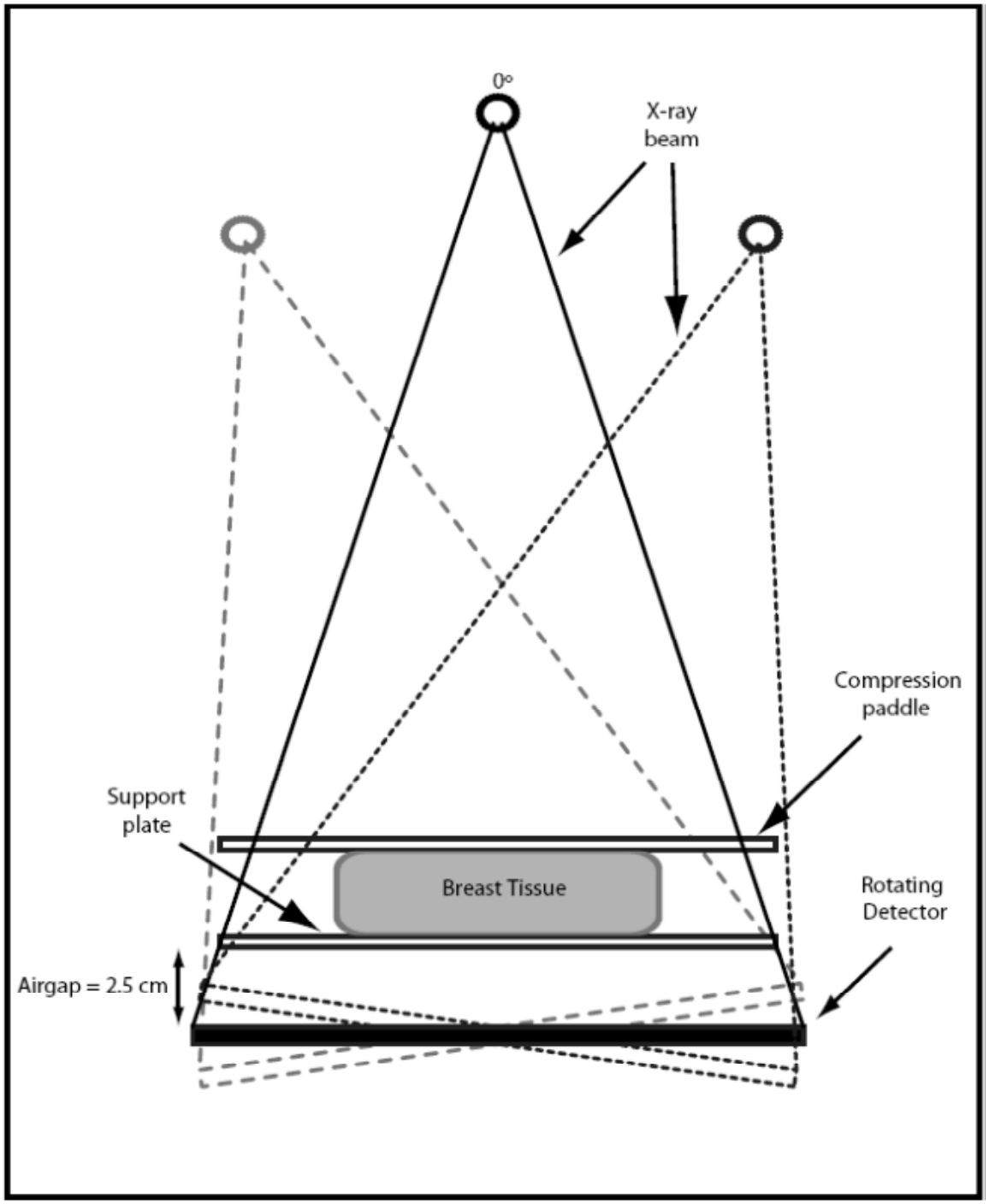

Figure 3.1– The Hologic Selenia Dimensions system acquires 15 projections, spaced evenly, over a 15 degree angular range with the center projection (tomosynthesis angle = 0 deg) perpendicular to the image detector plane. The source-to-imager distance (SID) is 70 cm, with a 2.5 cm airgap between detector and the breast support plate. The image detector measures 24 cm x 28 cm. (Image not to scale).

Under clinical conditions, in fully automated mode, the system determines the tube voltage, the tube current-exposure time product, and, for FFDM acquisitions only, the filter. To determine the X-ray tube voltage and, in FFDM mode only, the filter material, the system uses the thickness of the compressed breast. To determine the tube current-exposure time product, a single low-dose scout exposure is performed before image acquisition and the signal at the detector is analyzed. For a DBT acquisition, this scout exposure is performed with the tube in the -7.5 degrees position.

3.2.1 Automatic Exposure Control Selections

In order to determine the range of X-ray tube voltages and tube current-exposure time products the AEC would choose when imaging breasts of varying thicknesses and compositions, compressible phantoms of distilled water/olive oil mixtures in malleable plastic containers were made and imaged in the cranio-caudal (CC) view. These materials were chosen since water and olive oil are good analogs for breast glandular and adipose tissues, respectively.[73, 74] Both DBT and FFDM acquisitions were performed for each phantom at least three times, and the observed settings were averaged. These phantoms were made to simulate clinically encountered breasts, and ranged from 2 cm to 8 cm thick, in 1 cm steps, when compressed. The thickness was determined by both checking the system readout and measuring the separation between the compression paddle and the support plate using calipers. The glandular fraction of these phantoms was fixed at 0% (100% olive oil), 14.3%, 25%, 50%, 75%, and 100% (100% water). The 14.3% glandular compressible phantom was included as it has been recently shown by Yaffe et al[75] that this is the glandular fraction of an average breast. In addition, while the study by Yaffe et al[75] also found that patient breasts rarely have glandular fractions above 45%, phantoms with glandularity up to 100% were included to allow for comparison with previously published studies. Acquisitions were also performed with a commercial 50% glandular mammography phantom (Model 082, CIRS Inc., Norfolk, VA), and the AEC chosen

settings compared to those recorded for the 50% glandular compressible oil/water phantoms of selected thicknesses. The X-ray tube voltage, filter selection, and tube current-exposure time product chosen for each phantom was recorded, and are heretofore referred to as a designated X-ray spectrum. The inclusion of the filter section was necessary as the system selects either rhodium or silver as added filtration in the X-ray tube when performing FFDM acquisitions and aluminum as added filtration when performing DBT acquisitions.

3.2.2 X-Ray Tube Output Characterization

The X-ray source of the tested system uses a tungsten target and either a 50 μm thick rhodium or silver filter for FFDM acquisitions and a 0.7 mm thick aluminum filter for DBT acquisitions. The X-ray spectra can vary between 20 kVp and 49 kVp, as either selected by the AEC or manually.[72]

Each of the designated X-ray spectra, a combination of added filtration selection and tube voltage, observed to have been chosen for either a DBT or FFDM acquisition of the phantoms was modeled according to the method described by Boone et al,[57] and matched to the first half value layer (HVL) determined empirically by measuring the X-ray tube output using a calibrated ACCU-DOSE dosimeter (Radcal Corp, Monrovia, CA) and a dedicated mammography ionization chamber (Model 10X6-6M, Radcal Corp, Monrovia, CA). The ionization chamber was placed on the breast support plate at the central ray of the zero angle projection, which is the location for the reference air kerma used by the Monte Carlo simulations, and the air kerma measurement was recorded for a single acquisition with the tube current-exposure time product set at 100 mAs. Subsequently, additional thicknesses of aluminum (99.5% minimum, Model Number 1901017, Unfors Instruments, Inc., Hopkinton, MA) were added at the output port of the X-ray tube, and the exposure was measured again with each thickness. For accuracy in these and all subsequent measurements, three measurements were performed for each

thickness of aluminum and the results averaged. The HVL of the X-ray spectral models was matched to the measured HVL by varying the thickness of the modeled filter of the X-ray tube. The expected air kerma for each additional thickness of aluminum was calculated from these models and compared to that measured by the dosimeter. In addition, to determine the relationships between the tube current-exposure time product and air kerma, measurements were also taken for each X-ray spectrum using at least three different tube current- exposure time product settings.

All measurements for the tomosynthesis X-ray spectra were performed with the X-ray tube locked in the central projection position for all 15 projections, since this is the definition of the reference air kerma used by this study's Monte Carlo simulation to normalize its glandular dose output. The reference air kerma was defined in this manner to simplify its measurement, avoiding the need to locate the central ray for each tomosynthesis projections.[19]

3.2.3 Monte Carlo Simulations

A Monte Carlo simulation was developed and implemented in C++, based on the Geant4 Monte Carlo simulation toolkit,[54, 55] similar to those described in Sechopoulos et al.[19-21] In this study, only the FFDM and DBT cranio-caudal (CC) views were studied. The simulation also includes the breast to be imaged, with the previously described varying glandular fractions, a chest to nipple distance of 10 cm, a skin thickness of 0.4 cm, and compressed to thicknesses matching the compressible phantoms. Breasts of a single chest-to-nipple distance were simulated since previous studies have shown that both the normalized glandular dose for the central projection, and the mean of the relative glandular dose (RGD) for a complete symmetric tomosynthesis acquisition, especially in the CC view and for the limited angular range used by this system (± 7.5°), do not vary substantially with breast size.[19] In addition, the AEC selected the same acquisition protocols for breast phantoms of the same thickness and composition but different size.

Within each simulation, 1 million (10^6) X-rays of the same energy, from 5 keV to the maximum energy level of the spectrum selected by the AEC, in 0.5 keV steps, are emitted and tracked, and each energy deposition in the breast glandular tissue is recorded, according to the method of Boone et al[76] and Wilkinson and Heggie.[77] The monochromatic results are then combined as described by Boone et al[78] using the developed spectrum models to obtain the normalized mean glandular dose (both D_gN_{FFDM} and D_gN_{DBT}) in units of milliGray per milliGray air kerma.

The previously determined relationships between the tube current-exposure time product and air kerma and the recorded imaging techniques chosen by the AEC were used to calculate the total air kerma used for each size and glandularity of breast imaged. Finally, this was combined with D_gN_{FFDM} and D_gN_{DBT} to calculate the mean glandular dose to the breast for each imaging modality.

3.3 Results

3.3.1 Automatic Exposure Control Selections

The imaging techniques, comprising X-ray tube voltage and tube current-exposure time product, chosen by the AEC for each compressible phantom are shown in Table 3.1 and Table 3.2 for DBT and FFDM, respectively. A total of 14 distinct X-ray spectra were observed to have been chosen by the AEC, one for each compressed phantom thickness for each imaging modality.

In most cases, the AEC called for higher tube current-exposure time product settings with increasing phantom thickness and increasing glandular fraction, as expected. However, certain observed measurements do not fit these trends. Those for the DBT acquisitions of the 8.0 cm thick phantoms with glandular fractions of 1%, 14.3%, 25%, and 50% were lower than those of the 7.0 cm thick phantoms. That for the 3.0 cm phantoms with 100% glandular fraction was lower than those for the phantom of the

same thickness with 75% glandular fraction and for the 100% glandular 2.0 cm thick phantom. These inconsistencies remained even with repeated measurement. It was verified that the system readout for compressed breast thickness agreed with the caliper measurements in all cases.

Table 3.1- AEC SELECTIONS for DBT – The X-ray Tube Voltage and Tube Current-Exposure Time Product Settings Determined By the Automatic Exposure Control for Digital Breast Tomosynthesis. Note that the values shown here are the total tube current-exposure time product for all 15 projections. For DBT, this system always uses an aluminum filter.

Table 3.1 Tube Current- Exposure Time Product (mAs)

Breast Thickness (cm)	Tube Voltage (kVp)	Glandular Fraction (%)					
		1	14.3	25	50	75	100
2	26	31.0	31.0	31.7	31.3	33.0	46.3
3	28	33.0	33.0	33.0	34.0	37.7	35.7
4	29	45.0	45.0	45.0	45.0	49.0	65.3
5	31	51.0	51.0	51.0	52.0	66.0	73.0
6	33	60.0	60.0	60.0	60.0	71.7	87.0
7	35	72.0	72.0	72.0	72.0	87.3	106.3
8	38	69.0	69.0	69.0	69.0	89.0	111.7

Table 3.2 - AEC SELECTIONS for FFDM – The X-ray Tube Voltage and Tube Current-Exposure Time Product Settings Determined By the Automatic Exposure Control for Full Field Digital Mammography

Table 3.2 Tube Current-Exposure Time Product (mAs)

Breast Thickness (cm)	Tube Voltage (kVp)	Filter (Rh/Ag)	Glandular Fraction (%)					
			1	14.3	25	50	75	100
2	25	Rh	29.0	30.3	38.0	39.3	44.0	60.3
3	26	Rh	32.0	36.0	52.0	58.0	67.0	65.3
4	28	Rh	33.7	42.7	49.0	71.3	84.3	117.3
5	29	Rh	51.7	64.0	79.0	114.0	170.7	200.3
6	31	Rh	78.3	89.0	90.7	150.3	216.0	315.0
7	30	Ag	78.7	101.0	128.0	159.0	269.3	383.0
8	32	Ag	88.0	102.0	128.7	164.3	299.7	468.7

Settings chosen by the AEC for the 50% glandular compressible phantoms matched closely those chosen for the Model 082 50% glandular commercial mammography phantom for thicknesses of 2.0 cm, 4.0 cm, 6.0 cm, and 8.0 cm for both the DBT and FFDM modes (Table 3.3).

Table 3.3 – AEC Selections for Phantoms – Comparison of the Tube Current-Exposure Time Product Settings for Digital Breast Tomosynthesis and Full Field Digital Mammography between the CIRS Model 082 50% Glandular Phantom and the 50% Glandular Compressible Phantom. Note that the X-ray spectra selections by the AEC were identical, in terms of tube voltage, for phantoms of the same thickness.

Table 3.3 Tube Current- Exposure Time Product (mAs)				
Imaging Modality (DBT/FFDM)	Breast Thickness (cm)	Tube Voltage (kVp)	CIRS Model 082 Phantom (mAs)	50 % Glandular Compressible Phantom (mAs)
DBT	2	26	32.0	31.3
	4	29	45.0	45.0
	6	33	60.33	60.0
	8	38	69.33	69.0
FFDM	2	25	39.67	39.33
	4	28	71.67	71.33
	6	31	151.0	150.33
	8	32	166.33	164.33

3.3.2 X-ray Tube Output Characterization

The calculated expected change in air kerma of the developed X-ray spectrum models deviated by less than 4% from the measured air kerma changes recorded with the dosimeter when additional thicknesses of aluminum were placed under the X-ray tube output port. This comparison is shown for two example spectra: a 30 kVp W/Al tube/filter combination selected for DBT, and a 25 kVp W/Rh tube/filter combination selected for FFDM in Figure 3.2. The relationships between air kerma and tube current-exposure time product were found to be linear for all the designated X-ray spectra chosen by the AEC system with R^2 values greater than 0.99).

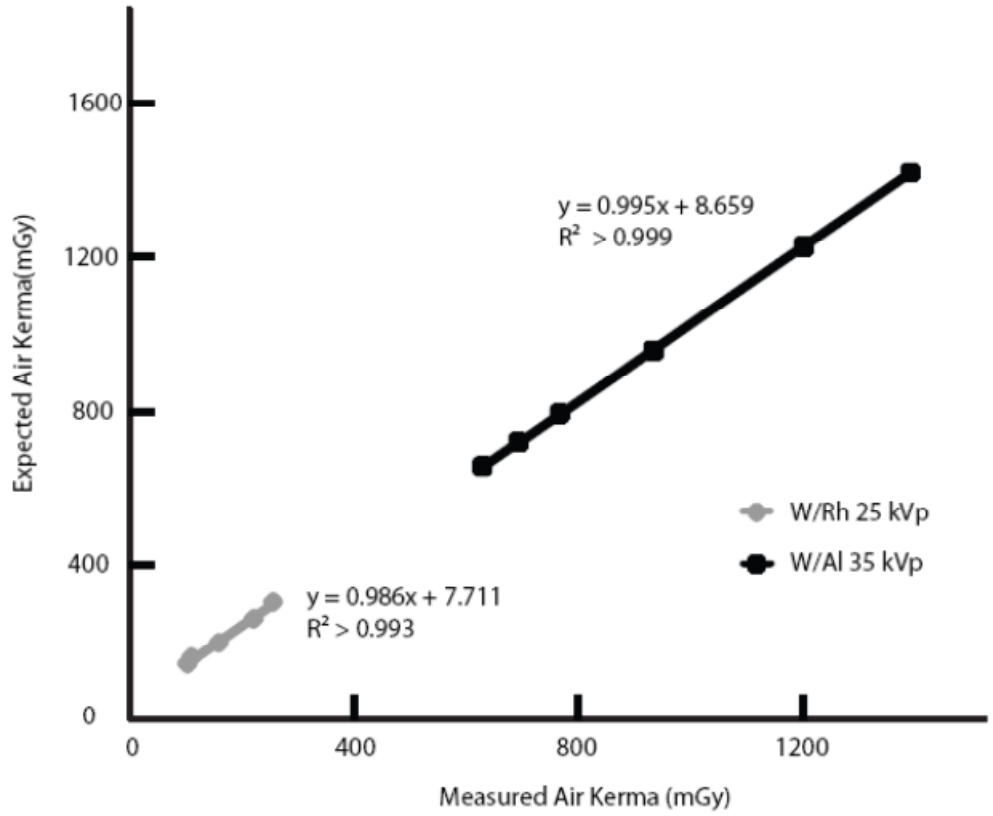

Figure 3.2 – X-ray Spectra Models – Two of the developed X-ray spectra models are shown here: a 35 kVp W/Al tube/filter combination selected for DBT, and a 25 kVp W/Rh tube/filter combination selected for FFDM. The calculated estimated air kerma agreed well with the measured air kerma ($R^2 > 0.99$).

3.3.3 Monte Carlo Simulations

The results for the normalized mean glandular dose for DBT (D_gN_{DBT}) and for FFDM (D_gN_{FFDM}) are shown in Table 3.4 and Table 3.5 respectively. As can be expected, both D_gN_{DBT} and D_gN_{FFDM} decrease with increasing glandularity and with increasing breast thickness in general. There is one exception of note, the phantoms measuring 8.0 cm thick exhibit a D_gN_{DBT} greater than that of the 7.0 cm thick phantoms of the same glandular fraction. However this can be attributed to the different X-ray spectra chosen by the AEC for phantoms of differing thicknesses.

Table 3.4 - D_gN_{DBT} MC Results

Table 3.4 Normalized Mean Glandular Dose (mGy/mGy air kerma)

Breast Thickness (cm)	Glandular Fraction (%)					
	1	14.3	25	50	75	100
2	4.05	3.90	3.78	3.52	3.28	3.06
3	3.44	3.27	3.14	2.87	2.64	2.42
4	3.25	3.08	2.95	2.68	2.44	2.22
5	3.05	2.88	2.75	2.48	2.25	2.05
6	2.89	2.72	2.60	2.34	2.12	1.92
7	2.75	2.59	2.47	2.22	2.01	1.82
8	2.79	2.63	2.51	2.26	2.04	1.85

Table 3.5 - D_gN_{FFDM} Monte Carlo Results

Table 3.5 Normalized Mean Glandular Dose (mGy/mGy air kerma)

Breast Thickness (cm)	Glandular Fraction (%)					
	1	14.3	25	50	75	100
2	4.23	4.07	3.95	3.68	3.42	3.19
3	3.89	4.07	3.56	3.25	2.98	2.73
4	3.46	3.27	3.12	2.81	2.55	2.31
5	3.15	2.95	2.81	2.51	2.26	2.03
6	2.87	2.68	2.55	2.26	2.02	1.81
7	2.66	2.49	2.36	2.09	1.87	1.68
8	2.54	2.37	2.24	1.98	1.77	1.59

Calculations of the total mean glandular dose for both DBT (MGD_{DBT}) and FFDM (MGD_{FFDM}), are shown in Table 3.6 and Table 3.7 respectively. Both MGD_{DBT} and MGD_{FFDM} generally decrease with increasing glandularity and increase with breast thickness (Figure 3.3).

Table 3.6 - MGD_{DBT} Calculated Results

Table 3.6 MGD_{DBT} (mGy)						
	Glandular Fraction (%)					
Breast Thickness (cm)	1	14.3	25	50	75	100
2	0.764	0.735	0.727	0.670	0.657	0.857
3	0.813	0.774	0.744	0.703	0.721	0.624
4	1.21	1.14	1.10	0.994	0.989	1.22
5	1.56	1.48	1.41	1.30	1.51	1.52
6	2.07	2.12	2.18	2.10	2.43	2.85
7	2.76	2.60	2.48	2.23	2.45	2.71
8	3.26	3.07	2.93	2.64	3.08	3.52

Table 3.7 - MGD_{FFDM} Calculated Results

Table 3.7 MGD_{FFDM} (mGy)						
	Glandular Fraction (%)					
Breast Thickness (cm)	1	14.3	25	50	75	100
2	0.309	0.313	0.389	0.376	0.395	0.516
3	0.392	0.461	0.581	0.592	0.627	0.560
4	0.455	0.544	0.595	0.779	0.832	1.05
5	0.660	0.775	0.920	1.20	1.63	1.73
6	1.09	1.16	1.12	1.68	2.17	2.85
7	1.22	1.48	1.79	1.99	3.03	3.88
8	1.53	1.66	2.00	2.28	3.74	5.26

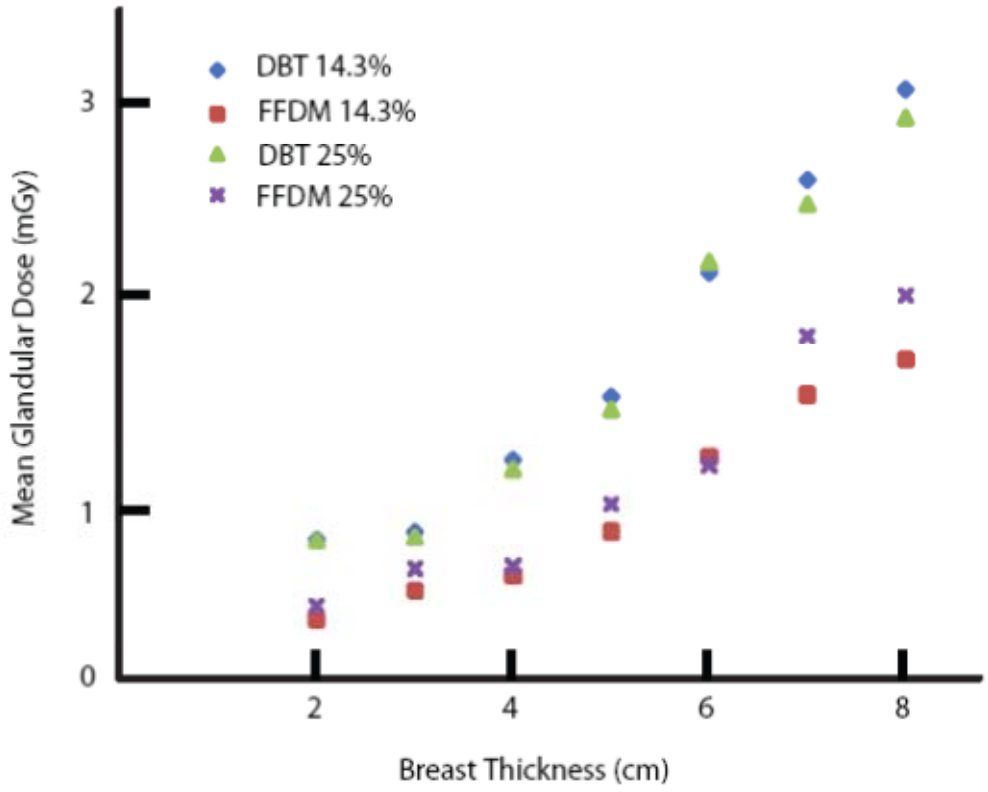

Figure 3.3 – Mean Glandular Dose vs Breast Thickness – Both MGD_{DBT} and MGD_{FFDM} increased with thickness and with glandular fraction for most breasts.

The ratio of MGD_{DBT} to MGD_{FFDM} is shown in Table 3.8. For most of the breast phantom sizes in this study, a single-view DBT acquisition results in a mean glandular dose of less than 2 times that of a single-view FFDM acquisition. However, there are notable exceptions, including two breasts of 14.3% glandular fraction: those of thickness 2.0 and 4.0 cm.

Table 3.8 – Ratio of MGD_{DBT} (from Table 3.7) to MGD_{FFDM} (from Table 3.6)

Table 3.8 MGD_{DBT}/MGD_{FFDM}

Breast Thickness (cm)	Glandular Fraction (%)					
	1	14.3	25	50	75	100
2	2.45	2.35	1.87	1.76	1.65	1.65
3	2.08	1.67	1.28	1.19	1.14	1.11
4	2.63	2.11	1.86	1.27	1.19	1.16
5	2.36	1.88	1.53	1.08	0.930	0.880
6	1.90	1.83	1.95	1.25	1.12	1.00
7	2.26	1.76	1.39	1.12	0.810	0.700
8	2.13	1.85	1.47	1.16	0.820	0.670

3.4 Discussion

The results of this study reveal that for the breast phantoms representing the most commonly clinically encountered breasts (2.0-8.0 cm, 1%-50% glandular fraction), a CC view acquisition can result in a mean glandular dose of 0.309 mGy to 2.28 mGy for FFDM and of 0.670 mGy to 3.26 mGy for DBT. Furthermore, for the breast phantom representing the "average" breast (5.0 cm thick, 50% glandular fraction),[79, 80] the mean glandular dose for the FFDM and DBT acquisitions are 1.20 mGy and 1.30 mGy, respectively; resulting in a difference of only 8% between the two modalities, and a fusion 2D/3D imaging study results in a mean glandular dose of 2.499 mGy. Hendrick et al. reported that the mean glandular dose for the 5 cm to 6 cm thick breasts included in the American College of Radiology Imaging Network's Digital Mammographic Imaging Screen Trial (DMIST)[81] ranged from 2.06 mGy to 3.01 mGy.[81] Therefore, the Hologic Selenia Dimensions system, in either FFDM or DBT mode results in a considerably lower mean glandular dose than those used in DMIST, and the combined FFDM/DBT study results in a dose comparable to that of a FFDM only acquisition in DMIST. For the breast phantom that represents the average breast from a recent study (thickness = 6.0 cm, glandular fraction = 14.3%),[75] a CC view acquisition results in a mean glandular dose of 1.16 mGy for FFDM and 2.12 mGy for DBT, a difference of 83%. Thus, a combined

FFDM/DBT breast imaging study of this breast using the Selenia Dimensions system results in a mean glandular dose of 3.28 mGy, 2.8 times that of a single-view FFDM and slightly higher than the FFDM dose range reported in DMIST for the 5 cm to 6 cm thick breasts.

For some of the thicker, denser (i.e. a glandular fraction of 75% and 100%), breasts, a DBT acquisition will actually result in a lower mean glandular dose – as low as 67% as much radiation dose as a FFDM acquisition. However, this is a result of the fact that the tube current-exposure time product chosen by the AEC for a FFDM acquisition of these clinically improbable breasts is very high, ranging from 269 mAs to 469 mAs, resulting in a much higher MGD_{FFDM}.

While previous studies have explored the normalized dosimetry of digital breast tomosynthesis and its comparison to digital mammography,[19, 21] providing dose conversion coefficients in units of dose per reference exposure/air kerma, this study provides a detailed characterization of the dosimetry in absolute terms, including taking into consideration the acquisition protocols automatically selected by the AEC for different breasts. In addition, the combination of measurements of the system's X-ray tube air kerma output, observation of the system's automated selections of X-ray spectra and tube current time-exposure time products for breast phantoms of varying sizes and tissue compositions, and system-specific Monte Carlo simulations to determine the resulting dose to the patient distinguish this work from those studies. To the best of our knowledge, this approach has not been previously used to characterize the dosimetry of any DBT system, although it has been used in the dosimetry of a dedicated breast Computed Tomography system.[74] It should also be noted that while in this experiment we are probing the AEC behavior, its exact size and location is not important, as long as the breast phantom is located in the standard mammography position for the compressed breast (central to the detector at the chest-wall edge).

Although this has yet to be proven possible, if in the future a single-view DBT acquisition replaces the current two-view FFDM screening exam, as has been discussed as a possibility,[25, 27, 82, 83] this would result in a mean glandular dose savings, as the former results in a mean glandular dose of less than 2 times that of a single-view FFDM acquisition in most clinically encountered breasts (Table 3.3). Furthermore, if screening with DBT results in a reduction in recall rate, as has been suggested by previously published studies,[25-28, 84] this could result in a reduction in population dose, since there would be a reduction in diagnostic work-up mammographic acquisitions, which often require acquisition of four to six additional views.

The results of this study are limited to providing the mean glandular dose delivered to homogeneous phantoms representing breasts of various glandular fractions. Future studies could include recording the AEC selections for both FFDM and DBT acquisitions of the same patient, followed by a comparison of the resulting mean glandular doses with an assumption of the glandular fraction of the patient breast. The use of homogeneous phantoms could also be a limitation, since in clinical conditions, the AEC settings are based on the signal towards the center of the detector, and therefore what is normally the densest part of the breast. Thus, a homogeneous phantom results in the AEC settings normally used for a breast with an overall lower glandular fraction; however, this would not affect the MGD_{DBT}/MGD_{FFDM} ratios found. In addition, this study investigated the resulting dose to the breast from CC view DBT and FFDM acquisitions and future studies of the resulting dose from medio-lateral oblique (MLO) images are needed. To perform this study for the MLO view, a realistic depiction of the portion of the pectoralis muscle present within the compressed tissue and in the field of view would need to be incorporated into the breast phantom to include its effect on the AEC's selected acquisition protocol. We are not aware of any reports on how much muscle tissue should be included in the different sized breast phantoms to represent clinically encountered conditions, nor of a reasonable material to mimic its attenuation

properties. Furthermore, although reports have included partial data on compressed thicknesses for the same breast in the CC and MLO views,[85, 86] a comprehensive study on this relationship still needs to be performed. Finally, it should be noted that mean glandular dose does not communicate the large variation in local dose deposition, which, as shown by Sechopoulos et al,[74] can vary from 15-400% of the mean glandular dose.

3.5 Conclusion

This dosimetric characterization of a clinical DBT system is an important contribution to the field of knowledge, and has major implications for patient care. In addition, comprehension of a dosimetry study and the resultant dose of this particular imaging system in both FFDM and DBT mode are critical objectives of this thesis, as the improvement of DBT image quality must be achieved without the compromise of radiation dose.

CHAPTER 4

OBJECTIVE MODELS OF COMPRESSED BREAST SHAPES UNDERGOING MAMMOGRAPHY

In this chapter I present models of compressed breasts undergoing mammography based on objective analysis, that are capable of accurately representing breast shapes in acquired clinical images and generating new, clinically realistic shapes.

An automated edge detection algorithm was used to catalogue the breast shapes of clinically acquired CC and MLO view mammograms from a large database of digital mammography images. Principal component analysis (PCA) was performed on these shapes to reduce the information contained within the shapes to a small number of linearly independent variables. The breast shape models, one of each view, were developed from the identified principal components, and their ability to reproduce the shape of breasts from an independent set of mammograms not used in the PCA, was assessed both visually and quantitatively.

The PCA breast shape models of the CC and MLO mammographic views were found to be able to both reproduce breast shapes with strong fidelity and generate new clinically realistic shapes. Visual examination of modeled breast shapes confirmed these results. The models were also used to generate CC and MLO view mammogram breast shapes, using the mean PCA parameter values of these distributions and randomly generated values based on the fitted Gaussian distributions, which resemble clinically encountered breasts.

The PCA models of breast shapes in both mammographic views successfully reproduce analyzed breast shapes and generate new clinically relevant shapes. This work can aid in research applications which incorporate breast shape modeling, such as X-ray scatter correction, dosimetry, and image registration.

4.1 Significance

According to the American Cancer Society, over 287,000 new cases of breast cancer were estimated to be diagnosed in 2011.[87] Currently, screening mammography is regarded as the most effective tool for early breast cancer detection and diagnosis. Standard mammography is a low-cost, fast, non-invasive X-ray study that involves relatively low doses of ionizing radiation, but has room for improvement in performance, demonstrating a sensitivity and specificity of 83.5% and 90.9%, respectively.[88]

A large amount of the continued research being performed to improve the clinical performance of breast imaging requires the simulation of realistic patient breast models, and this has become an entire research area in itself.[89-96] Research on breast imaging dosimetry,[19, 21, 71, 76, 78, 97-101], lesion detectability,[102-104], technique optimization,[105, 106] image registration,[107] image texture[108] and X-ray scatter correction[109] has been performed using software breast models, representing both uncompressed and compressed breasts. In most studies involving a model of the compressed breast, only the cranio-caudal (CC) view has been simulated, with the breast represented as a semi-circle or semi-ellipse. In a few studies, models of the breast compressed for the medio-lateral oblique (MLO) view have been used, but the models were subjectively derived.[19-21, 110, 111] Although these models have proved appropriate, the development of more sophisticated models that better represent not only the "average" breast shape but also attempt to include the span of clinically realistic breast shape variations would be useful in many of the areas of research listed above that use software breast models.

Here we present models for the breast shape when compressed for both the CC and MLO views of mammography that leverage principal component analysis (PCA) to accurately reproduce breast shapes and to generate new, clinically realistic shapes.

4.2 Materials and Methods

These models of mammogram breast shape were developed in three distinct steps. First we assembled a large database of clinically acquired cranio-caudal (CC) and medio-lateral oblique (MLO) view mammograms and systematically catalogued the breast shapes. Then we performed principal component analysis (PCA) on these shapes to condense the information down to a reduced number of linearly independent variables. And finally we tested the models' ability to estimate the shape of breasts from mammograms that were not used in the PCA.

4.2.1 Mammogram Database

We have an anonymized database of 477 paired CC and MLO view digital mammograms. This database was previously compiled in our institution and as the anonymization process is irreversible, no further IRB oversight is necessary. These were acquired using a Senographe 2000D digital mammography system (GE Healthcare, UK) that utilizes an X-ray tube with molybdenum and rhodium targets with added molybdenum and rhodium filtration. The system has a 19.14 x 22.94 cm indirect detector with a pixel pitch of 100 μm.

This database was divided into 377 images of each view to be used in the formation of the PCA model, and 100 images of each view to be used for testing.

4.2.2 Automated Edge Detection

Each mammogram was automatically thresholded to obtain a binary image, Y_b, separating tissue from open field.[112] The extraneous tissue, i.e. tissue not part of the breast under examination, was removed to isolate only the breast tissue in the image via the extraction of connected components.[112] This was accomplished by repeated iterations of image dilation with a dilation structure B and image intersection with Y_b:

$$Y_k = (Y_{k-1} \oplus B) \cap Y_b \qquad \text{(Equation 4.1)}$$

We chose the midpoint of the length along the chest wall of the mammogram as a starting point for a 5x5 pixel initial guess and a 10x10 structure was chosen for B. The extraction was considered complete when $Y_k = Y_{k-1}$. Using this isolated binary image of the breast, the edge was then located by using Roberts cross gradient operators, and edge thinning.[62, 63, 112] The edge is extended to the chest wall, if needed, to form a set of points that depict the breast shape in the mammogram. This set was then uniformly sampled to generate a subset of 100 (x,y) coordinate points, in terms of distance rather than pixel index, for use in the PCA. This is very similar to a process we previously used to register breast shapes for digital tomosynthesis images.[109] For the purposes of this study, the automation of the breast tissue isolation and edge detection allows for many images to be processed quickly without requiring user input and subjective analysis of the images.

To validate the automated edge detection, a radiologist enrolled in a mammography fellowship evaluated the generated edges of 100 CC view mammograms and 100 MLO view mammograms, all of which were randomly selected from the database. Each mammogram was shown side by side with a view of the generated edge overlaid onto the mammogram, allowing the radiologist to simultaneously view the edge and the original image. .

4.2.3 Principal Component Analysis

The subset of points that delineate the shape of the breast in the mammogram can be defined as 100 correlated observations which we can reduce using PCA to a smaller set of linearly independent variables that encapsulates most of the information contained within the datapoints.[113, 114] PCA has been used in innumerable applications, including in image registration.[115-117] To perform the analysis, we formed the vector $\boldsymbol{x_i}$ as follows for each image i:

$$\boldsymbol{x_i} = [x_1, x_2, x_3, \dots, x_{100}, y_1, y_2, y_3, \dots, y_{100}]^T \qquad \text{(Equation 4.2)}$$

where the (x,y) pairs are the pixel coordinates that define the breast shape, in cm. We then formed a square covariance matrix, cov(X), from the mean-centered $\boldsymbol{x}$ vectors:

$$cov(\boldsymbol{X}) = \frac{\sum_{i=1}^{N}(\boldsymbol{x_i} - \bar{\boldsymbol{x}})(\boldsymbol{x_i} - \bar{\boldsymbol{x}})^T}{N-1} \quad \text{(Equation 4.3)}$$

where N is the total number of images processed and $\bar{\boldsymbol{x}}$ is the mean of the $\boldsymbol{x}$ vectors. The principal components of the dataset are then represented by the eigenvectors $\mathbf{e_j}$, which can be obtained by solving the eigenvalue problem:

$$cov(\boldsymbol{X})\boldsymbol{e_j} = v_j\boldsymbol{e_j} \quad \text{(Equation 4.4)}$$

where the associated eigenvalues (v_j) also represent the variance of each principal component. For each view (CC and MLO), the n principal components with the greatest v_j were selected and their respective $\mathbf{e_j}$ assembled into a 100 x n principal component matrix **E**. This matrix can be used to describe, to a high degree of accuracy, any of the breast edges $\boldsymbol{x_i}$ that were used in the PCA and forms the basis of our breast shape model. The PCA parameter vector $\boldsymbol{r_i}$ for image i can be calculated as follows:

$$\boldsymbol{r_i} = (\boldsymbol{E^T E})^{-1}\boldsymbol{E^T}(\boldsymbol{x_i} - \bar{x}) \quad \text{(Equation 4.5)}$$

Initial validation of the model was done by visually comparing the original breast edge of several mammograms from the foundational database to that estimated from its PCA parameters in $\boldsymbol{r_i}$. An estimate of the original breast edge, $\tilde{\boldsymbol{x}}_{\boldsymbol{i}}$, was calculated as follows:

$$\tilde{\boldsymbol{x}}_{\boldsymbol{i}} = \bar{x} + \sum_{j=1}^{n}\boldsymbol{r_{i,j}}\ \boldsymbol{e_j} \quad \text{(Equation 4.6)}$$

where n is the number of principal components chosen for **E**.

Histograms of the first six PCA parameters obtained for each of the breast shapes in the foundational database were generated, and each one was fitted with a Gaussian distribution.

All PCA analysis was performed using MATLAB® R2011b (MathWorks, Natick, MA).

4.2.4 PCA Model Testing

To test the PCA models of breast shape, we utilized 100 CC view and 100 MLO view mammograms that were not included in the PCA. Each test mammogram *m* underwent automated edge detection as described above to generate a subset of points that depict a test edge $\boldsymbol{t_m}$. Using the view-specific principal component matrix, $\mathbf{E_{CC}}$ or $\mathbf{E_{MLO}}$, we calculated the PCA parameter vector, $\boldsymbol{r_m}$, of each test edge, and the modeled edge $\tilde{\boldsymbol{t}}_{\boldsymbol{m}}$ as follows:

$$\boldsymbol{r_m} = (\boldsymbol{E^T E})^{-1}\boldsymbol{E^T}(\boldsymbol{t_m} - \bar{x}) \qquad \text{(Equation 4.7)}$$

$$\tilde{\boldsymbol{t}}_{\boldsymbol{m}} = \bar{x} + \sum_{j=1}^{n} \boldsymbol{r_{m,j}}\ \boldsymbol{e_j} \qquad \text{(Equation 4.8)}$$

The accuracy of $\tilde{\boldsymbol{t}}_{\boldsymbol{m}}$ was assessed by calculating the average distance error (ADE) between it and the original edge, $\boldsymbol{r_m}$. ADE was calculated as follows:

$$\text{ADE} = \frac{(T_b \cup Y_b) - (T_b \cap Y_b)}{0.5(L_r + L_t)} \qquad \text{(Equation 4.9)}$$

where T_b is a binary image, depicting the PCA modeled breast shape, created by shape filling an image of $\tilde{\boldsymbol{t}}_{\boldsymbol{m}}$, Y_b is the binary image of the original mammogram, and L_r and L_t are the lengths of the modeled edge and the original edge, respectively. Thus, this metric is formed by dividing the area that lies in between the edges formed by the modeled and original breast shapes by the average of the two edge lengths (Figure 4.1). A paired t-test was performed on the ADE values of each tested image between the PCA models generated with differing number of principal components.

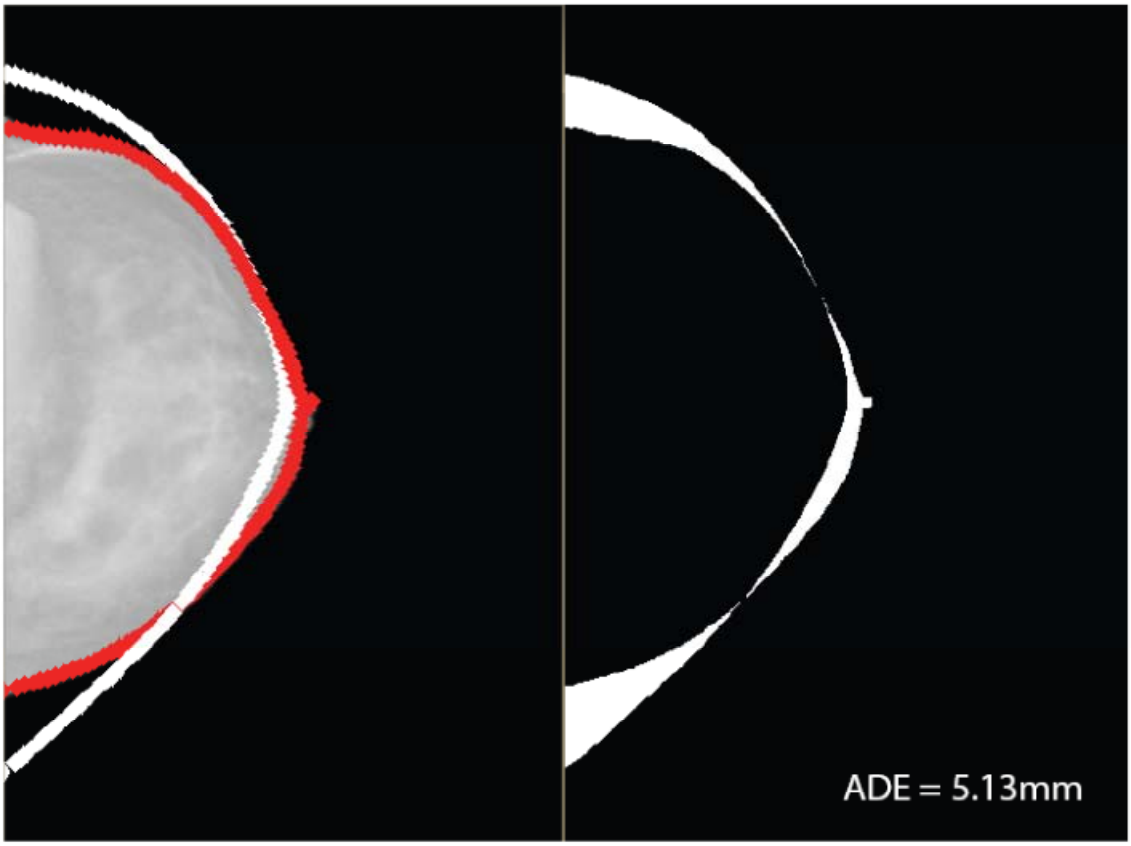

Figure 4.1 – Average Distance Error Calculation – The automatically generated breast edge (red) and the PCA modeled edge (white) are shown here on the left for a CC view mammogram. The image shown on the right represents $(\boldsymbol{T_b} \cup \boldsymbol{Y_b}) - (\boldsymbol{T_b} \cap \boldsymbol{Y_b})$, where T_b and Y_b are binary images depicting the PCA modeled breast shape and the original mammogram respectively. The area of the region in white shown in the right image is divided by the average of the length of the two edges to obtain the ADE.

4.2.5 PCA Parameter Characterization

To understand how the value of each PCA parameter affects the breast shape in the 6-component model, we generated novel shapes using selected values based on the mean and standard deviation of the fitted Gaussian probability distributions of the parameter histograms. For this, each parameter was varied individually, with the selection of five values ranging from its mean minus two standard deviations ($\mu - 2\sigma$) to $\mu + 2\sigma$ while the other parameters were held constant at their mean values. In this way, the effect of each parameter when its values span approximately 95% of the modeled clinically relevant range can be visualized.

4.2.6 PCA Parameters and Breast Size

We examined the relationship between breast size in the mammogram, both in terms of breast area and compressed breast thickness, and each of the calculated PCA parameter values from both $\boldsymbol{r}_m$ and $\boldsymbol{r}_i$. The thickness was retrieved from the mammogram database, and the breast shape area was calculated from the binary image, Y_b, of the original mammogram.

4.3 Results

4.3.1 Automated Edge Detection

Our automated edge detection algorithm proved very successful in isolating the breast tissue under examination from extraneous tissue present in the mammogram, and delineating a set of points that form the shape of the breast. Two examples each of CC view and MLO view breast shapes from the mammogram database (that used for the PCA) can be seen in Figure 4.2. The delineated breast edges are overlaid on the images in white. The radiologist enlisted in the validation study determined that all 200

automatically generated edges presented, evenly distributed between the two views, are successful representations of the breast shape edge.

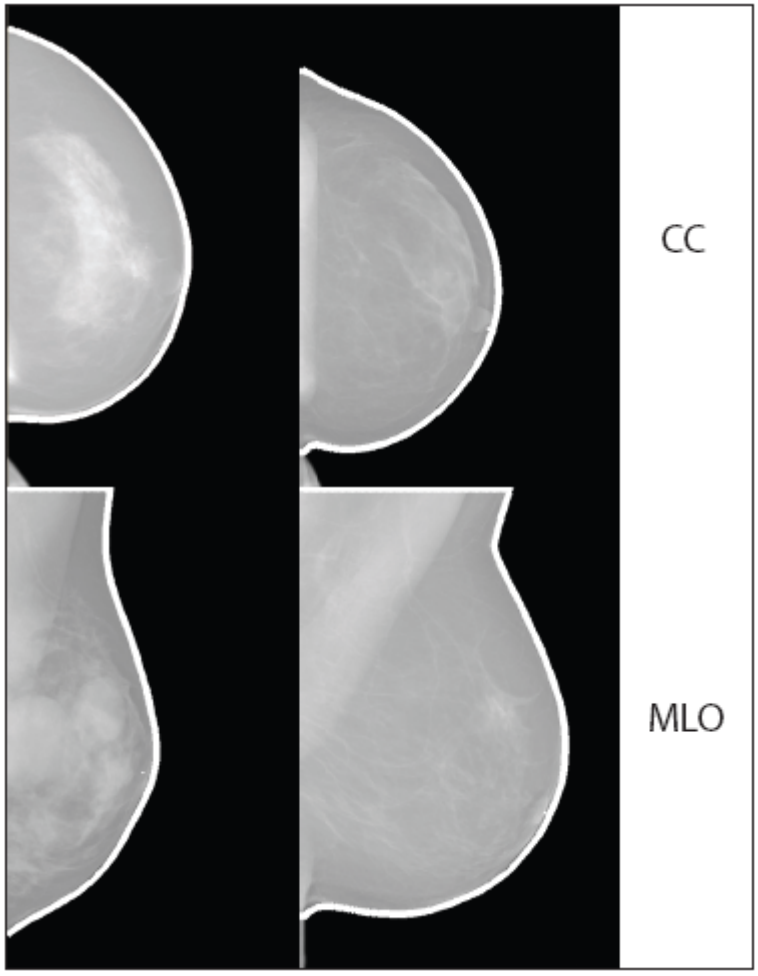

Figure 4.1 – Breast Edges – Shown here, in white, are edges generated by our automated edge detection algorithm on CC view (top) and MLO view mammograms (bottom). The extraneous tissue seen at the bottom of the CC view mammograms was excluded by our algorithm.

4.3.2 Principal Component Analysis

Our PCA of the CC and MLO view mammograms revealed that 93.3% and 85.4%, respectively, of the variance in the database images is contained in the first two

principal components, that 99.2% and 98.0% of the variance is contained in the first six principal components (Table 4.1). We have included the eigenvectors of the first six principal components of both models in the Supplemental Materials.

Table 4.1 – Percentage of total variance contained in the first six principal components – Shown here is the variance contained within each individual component and the cumulative variance contained within each component and its predecessors.

Table 4.1 Percentage of Total Variance

Principal Component	Cranio Caudal View/Cumulative (%)	Medio-lateral Oblique View/Cumulative (%)
α	74.8/74.8	62.3/62.3
β	18.5/93.3	23.1/85.4
γ	3.54/96.8	6.72/92.1
δ	1.03/97.9	3.13/95.2
ε	0.72/98.6	1.83/97.1
ζ	0.54/99.2	0.87/98.0

4.3.3 PCA Model Testing

The 2-component models produced reasonable recreations of breast shapes of mammograms from the testing set; three examples of breast shapes of each view are shown in Figure 4.3 (white outline), along with their ADEs. The upper left image depicts a CC-view mammogram for which the 2-component model generates a breast shape with a very low ADE that is similar to that of the 6-component model (0.88 mm to 0.79 mm, respectively). However, as can be seen in the upper right image, for some cases the 2-component model does not recreate the breast shape accurately, resulting in a higher ADE (7.15 mm). Visual inspection and the ADE values of the 4-component and 6-component PCA models for each view revealed that such models generate more accurate representations that include intricacies of some breast shapes. Test edges generated by the

6-component models can also be seen in Figure 4.3 (red outline). Some breast shapes are recaptured to a much greater degree by the 6-component model, such as the middle and right images shown in both rows in Figure 4.3. Table 4.2 shows the mean and maximum ADE for the breast shapes generated with the 2-, 4- and 6-component models for both the CC and MLO views.

Table 4.2 – Average Distance Error – The mean ADEs of the 2-component PCA model breast shapes was 2.99 mm for the CC view and 4.63 mm for the MLO view, but the 6-component model produced more accurate shapes on average and in the worst cases (maximum).

Table 4.2 Average Distance Error (ADE)

	2-Component Model (mm)		4-Component Model (mm)		6-Component Model (mm)	
View	Mean	Maximum	Mean	Maximum	Mean	Maximum
CC	2.99	7.80	1.47	3.92	0.90	1.94
MLO	4.63	13.6	2.14	6.25	1.43	4.49

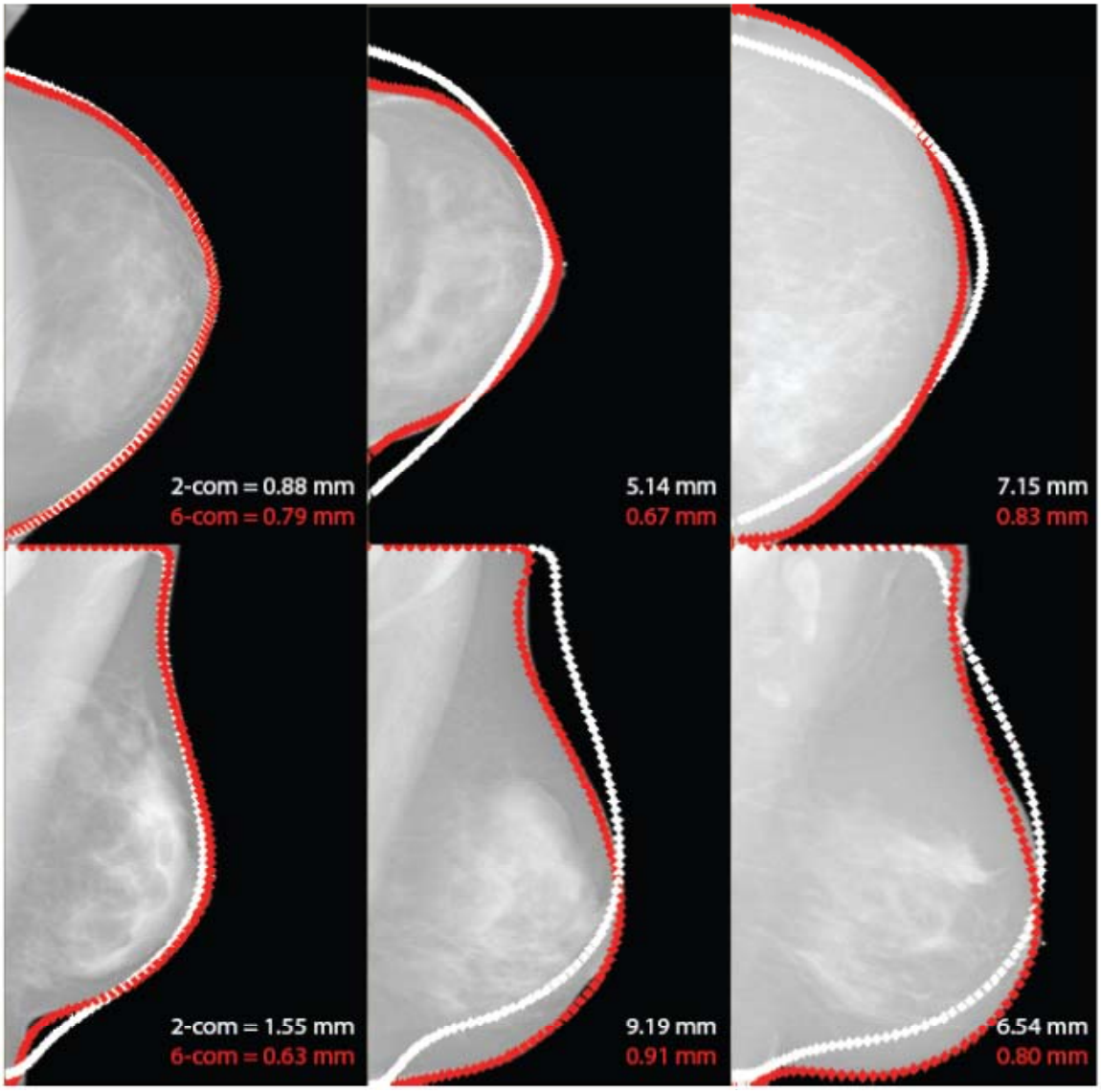

Figure 4.2 – Modeled breast shapes using 2 and 6 principal components – Shown here are breast shapes generated from 6-component PCA models (red) of the CC (top) and MLO (bottom) views overlaid on the shapes generated by 2-component models (white) and their respective ADEs. In some cases, the 2-component model generates adequate breast shape representations (left, top and bottom). However, in other cases (middle and right) a model with a greater number of components is needed to recapture the intricacies of the breast shapes.

We also compared the performance of the models with differing number of components by examining a box-and-whisker plot of the ADEs of each model (Figure 4.4). As can be seen, for every percentile mark (10^{th}, 25^{th}, median/50^{th}, 75^{th}, and 90^{th}),

models with more components exhibited lower ADEs than those of the same view with fewer components. Thus, as expected, models with more components consistently perform better. For each view, paired t-tests were performed between the ADE values for each PCA model, and all p-values were found to be less than 0.025.

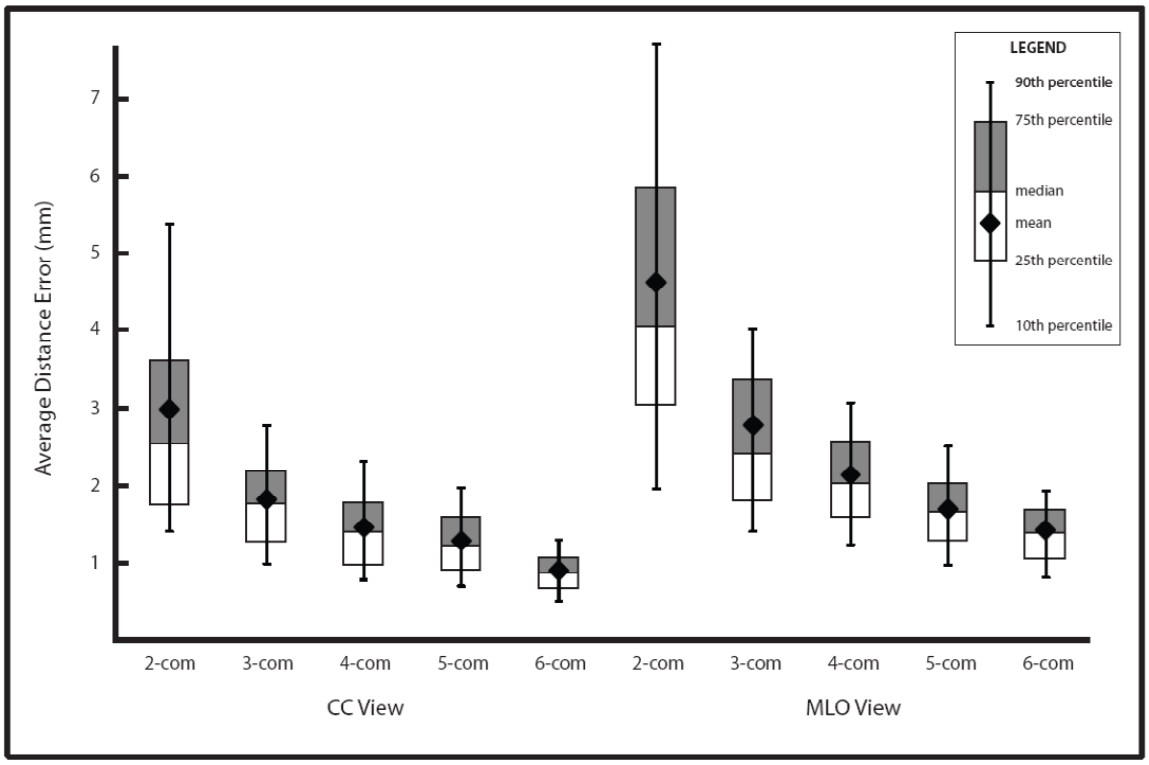

Figure 4.3 – Box-Whisker Plot of Average Distance Error – PCA models with more components consistently performed better than those of the same view with fewer components, exhibiting not only lower means but also lower ADEs for the 10^{th}, 25^{th}, 50^{th}/median, 75^{th}, and 90^{th} percentiles. Shown here are the models of the CC-view (left) and MLO-view (right) with components numbering from 2-6. All differences in paired ADE values of the PCA models with different number of principal components were statistically significant ($p<0.025$).

Histograms of the six PCA parameters for each view were fitted with Gaussian distributions and are shown in Figure 4.5, along with the mean and standard deviation of each fit.

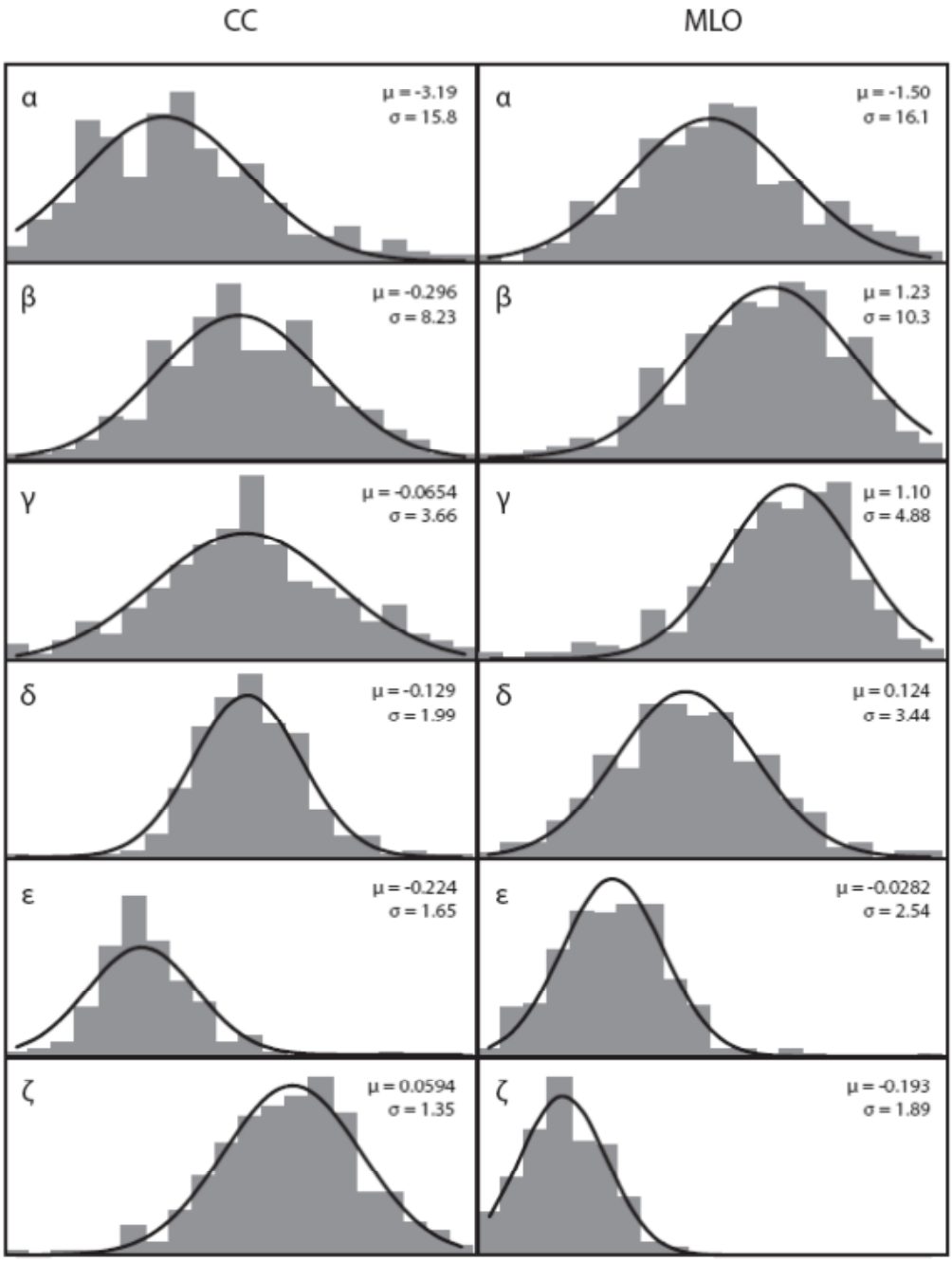

Figure 4.4 – PCA Parameter Distributions – The distributions of the first six PCA parameters (α, β, γ, δ, ε, ζ) of the PCA Breast Shape Models for both the CC (left) and MLO (right) views can be fitted with Gaussian distributions. The mean and standard deviation of each fit is also shown for each parameter.

4.3.4 Novel Breast Shape Generation

The 6-component objective PCA breast shape models presented here were also used to generate novel breast shapes, i.e. shapes that are clinically relevant but which were not derived from a clinical mammogram. This was achieved by using the view-specific principal component matrix, $\mathbf{E_{CC}}$ or $\mathbf{E_{MLO}}$, and selected values for the PCA parameter vector $\boldsymbol{r_m} = [\alpha, \beta, \gamma, \delta, \varepsilon, \zeta]$. First we generated the models' average breast shapes by using the mean values of the PCA parameters (Figure 4.6, top). It should be noted that this "average" breast shape in the CC view results in a shape that is similar to, but not quite, a semi-ellipse, with a chest wall-to-nipple distance of 10.1 cm and a distance along the chest wall of 20.2 cm, and a calculated area of 144 cm^2. This size corresponds well with the mean equivalent radius of 10 cm and the mean equivalent area of 157.3 cm^2 determined by Boone et al. from 82 patient CC mammograms.[80]

We also generated a random PCA parameter set, using the fitted Gaussian probability distributions of the parameter histograms, and generated a random clinically relevant breast shape for each view (Figure 4.6, bottom).

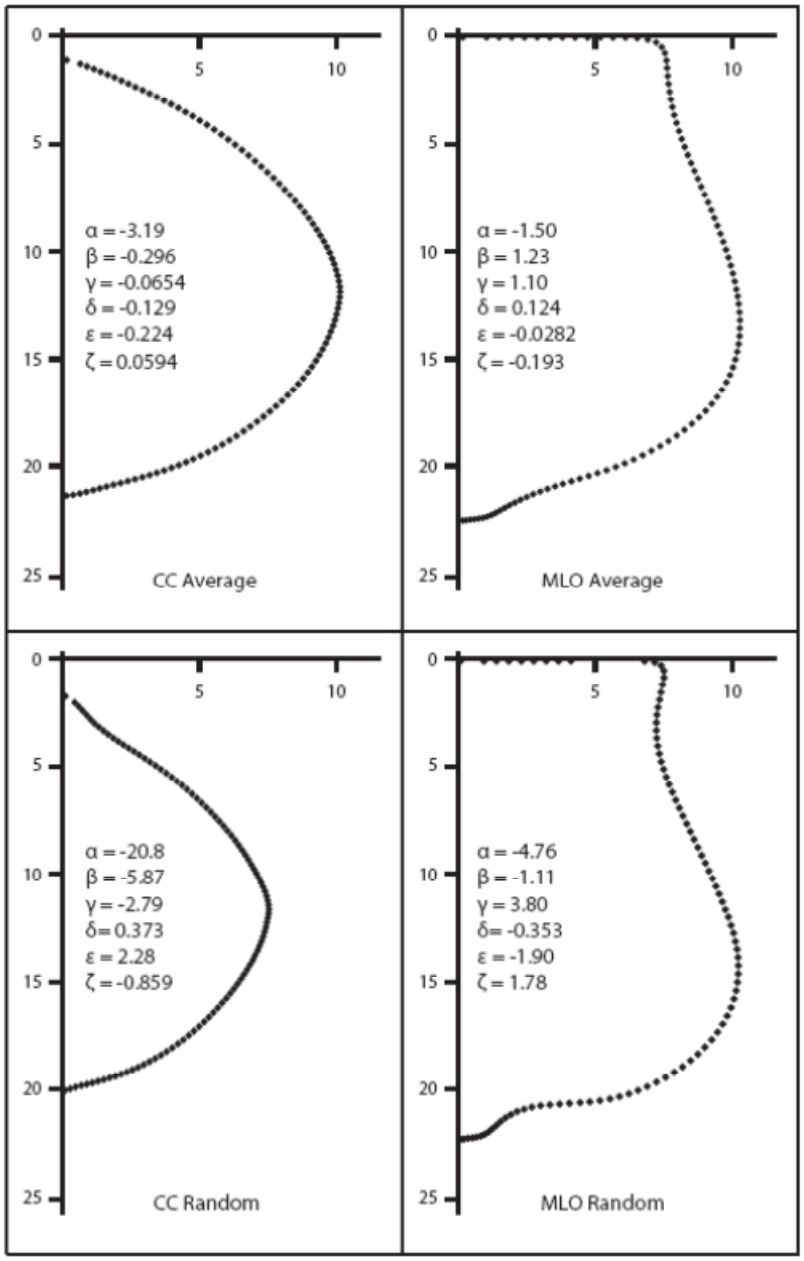

Figure 4.5 – Novel Breast Shapes Generated by the 6-component PCA Model – These breast shapes were generated by the selection of a set of values for the PCA data vector $\boldsymbol{r_m}$. The top CC and MLO view modeled breasts were generated using the mean values of the Gaussian fits for each principal component, so they represent an "average" breast shape. The bottom breast shapes were generated using randomly generated values from the Gaussian distributions for each principal component.

4.3.5 PCA Parameter Characterization

Figure 4.7 (CC view) and Figure 4.8 (MLO view) show how the breast shape model varies when the six PCA parameters are varied individually. As can be seen, the first PCA parameter, α, is the greatest determinant of the overall size and shape of the breast for both the CC and MLO view models, while the subsequent five parameters (β, γ, δ, ε, ζ) adjust different detailed shape features of the breast.

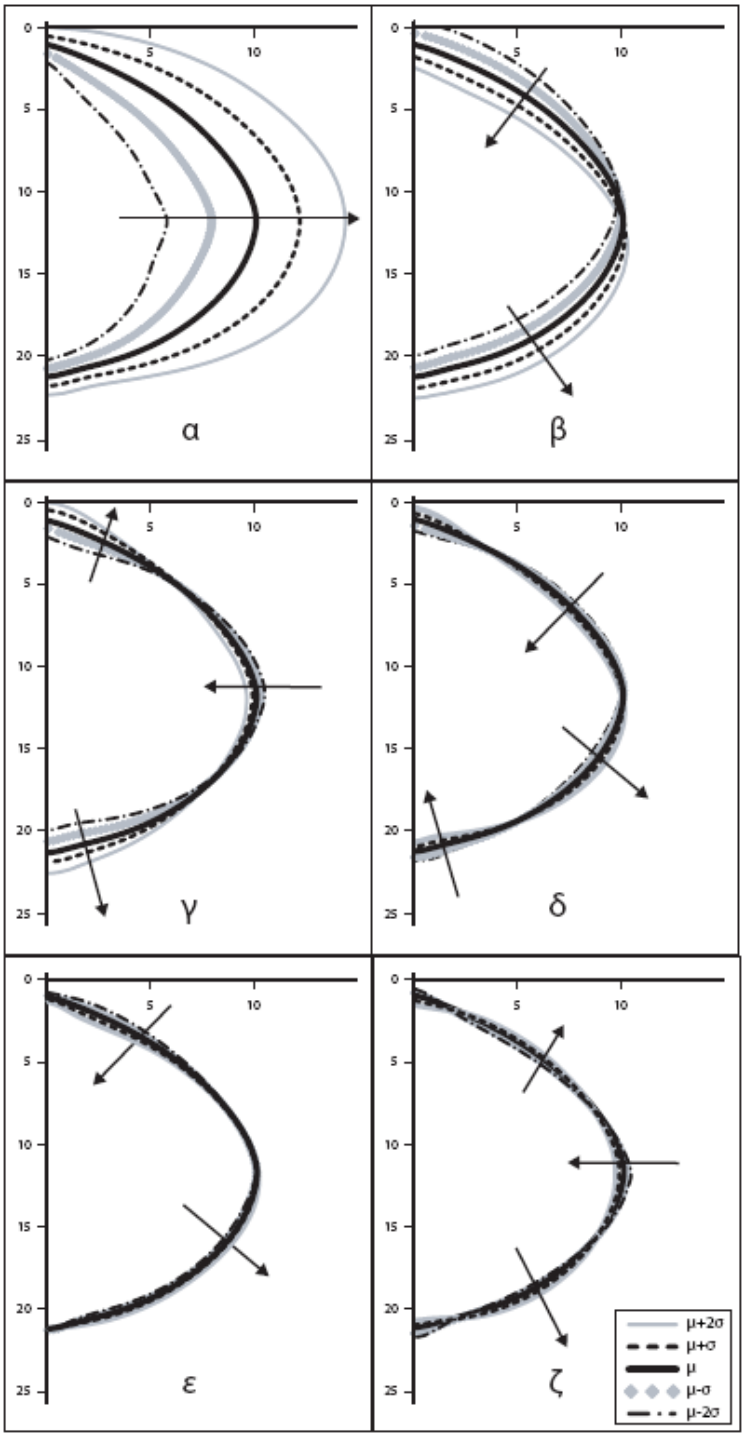

Figure 4.6 – Varying PCA parameters of the 6-component CC view model – Shown here are CC view breast shapes generated by varying each of the six PCA parameters (α, β, γ, δ, ε, ζ) individually. Each parameter was set to the mean (μ) value from the fitted Gaussian distributions shown in Figure 5, as well as μ +/- one standard deviation (σ), and μ +/- 2σ. The arrows indicate the overall direction in which that portion of the breast shape shifts as each parameter is increased. Within each panel, the average breast (in which all the values are set to their mean) is shown in solid black.

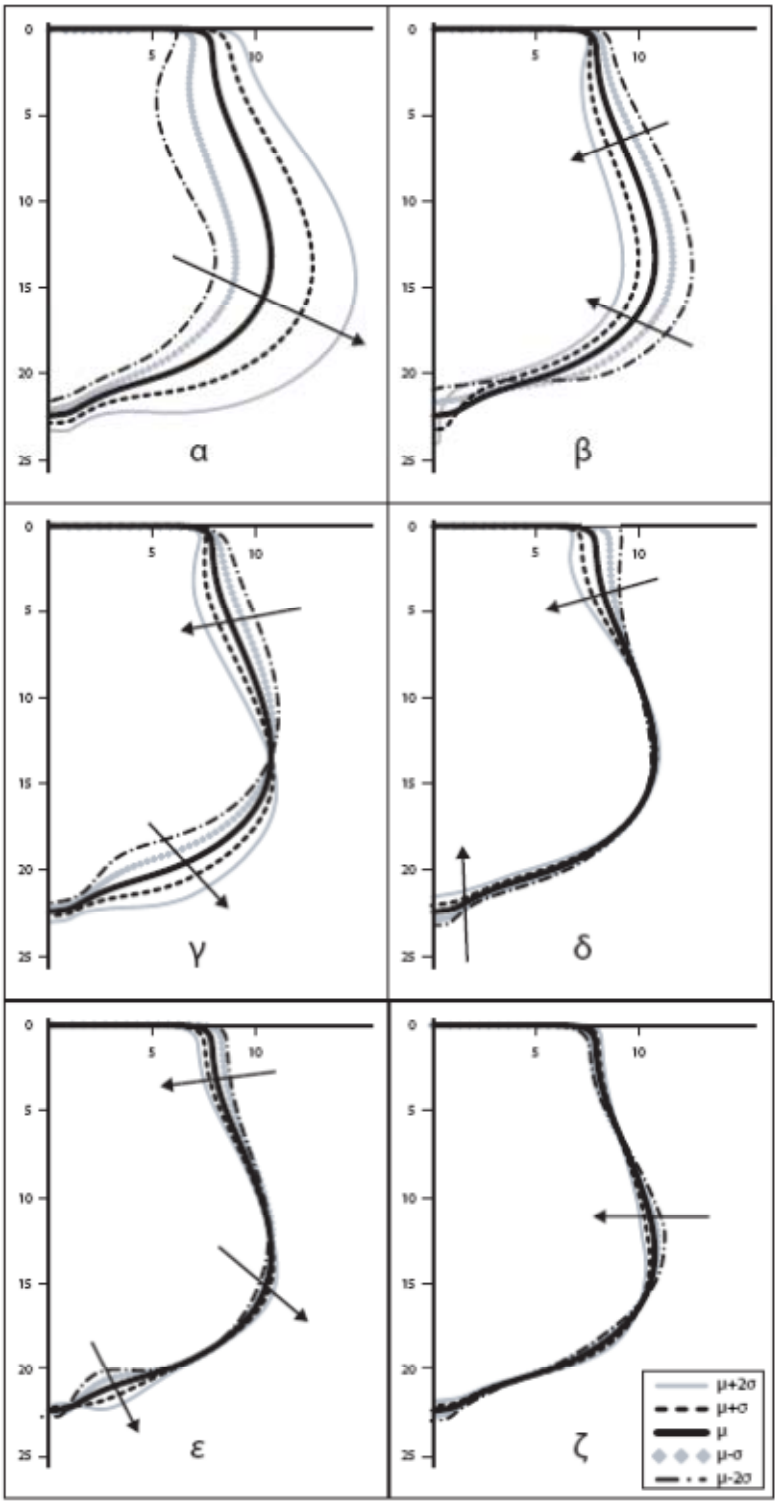

Figure 4.7 – Varying PCA parameters of the 6-component MLO view model - Shown here are MLO view breast shapes generated by varying each of the six PCA parameters (α, β, γ, δ, ε, ζ) individually. The process by which this was accomplished is the same as in Figure 7. The arrows indicate the overall direction in which that portion of the breast shape shifts as each parameter is increased. Within each panel, the average breast (in which all the values are set to their mean) is shown in solid black.

4.3.6 PCA Parameters and Breast Size

A strong relationship was found between the first PCA parameter, α, for both views and the breast area (Figure 4.9, left). However, there was no apparent relationship between the α parameter and the compressed breast thicknesses in either view (Figure 4.9, right). None of the other PCA parameters were found to correlate strongly with either measure of breast size in either view.

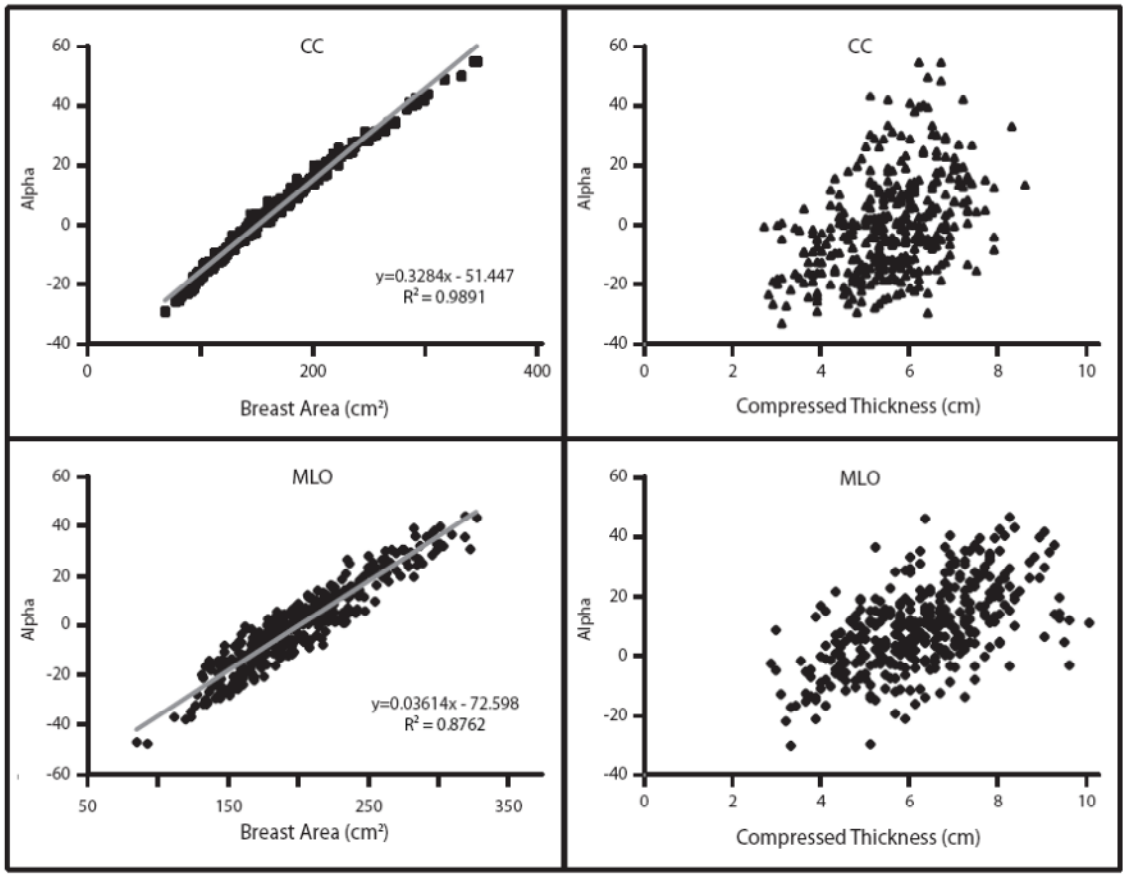

Figure 4.8 – Relationship Between α and Breast Size – For both the CC (top) and MLO (bottom) views, a strong relationship was found between α and the breast area (left). However, for the compressed breast thickness (right), there was no strong correlation found.

4.4 Discussion

In this study, we developed objective models of the compressed breast shape in the CC and MLO mammographic views using Principal Component Analysis (PCA) and tested the models' ability to both recreate breast shapes that were not part of the original analysis and to create novel clinically relevant breast shapes. Visual examination and quantitative analysis in the form of the average distance error (ADE) revealed that the 2-6 principal component models of breast shape presented here show great promise and could become useful tools for breast imaging applications, both for analyzing acquired images and for generating realistic breast shapes. However, the number of components to include in the PCA model depends on the application. In the generation of realistic breast shapes, the use of a high number of components should not be onerous and it would provide the greatest variability and detail in clinically-relevant shapes, so all 6 components could be included. However, for other applications it is possible that the inclusion of more components could introduce high complexity, making the use of a greater number of components unfeasible. For these cases, using only a 3- or 4-component model could still result in acceptable representations of the breast, given the low ADE values found (Table 4.2, Figure 4.4). In fact, the 2-component models generate a shape that resembles a breast undergoing mammography (Figure 4.3, white) in both the CC and MLO view, and recaptures the general curvature and closed edge of the breast, which might be sufficient for some applications.

Evaluation of each individual PCA parameter (Figure 4.7, 4.8) shows that the first PCA parameter, α, determines to a great extent the size of the breast shape in the modeled mammogram. We also found that α is strongly correlated with the breast shape area, but not compressed breast thickness. The relationships shown in Figure 4.9 can be a starting point for the user to generate new, clinically relevant breast shapes with control of the size of the generated breasts.

We have used the models presented here to characterize the breast shapes seen in almost 500 mammograms each of the CC and MLO views. As revealed by the ADE results (Table 4.2, Figure 4.4), our models have reduced the complexities of breast shapes to a small number of parameters; using just 2 to 6 values, we can describe to a great extent the breast shapes in the original mammograms. The use of the presented models to generate novel, clinically relevant breast shapes could aid research in 3D breast software phantoms by simplifying the mathematical representation of the breast shape. In addition, we have previously developed a software-based scatter correction algorithm for digital breast tomosynthesis imaging (DBT) that utilizes Monte Carlo X-ray simulations of standardized breast shapes, so the PCA models developed here facilitate the scatter correction algorithm's application to clinically encountered breasts.[109] We have performed preliminary testing with the central angle X-ray projection images of DBT imaging, and results have also been promising but a larger database of clinical images is required for further testing. In addition, this model forms the basis for a PCA based breast volume model, detailed in chapter 5, to be used for DBT simulations. And lastly, dosimetric characterization of breast imaging systems that are also based on Monte Carlo simulations could leverage our PCA model's ability to randomly generate breasts of different clinically relevant shapes quickly.[19, 74, 118]

It is also important to note that the PCA models developed here, and the methodology we have employed, can be used independently of our automated edge detection algorithm. Any user-defined breast edge representation can be used in conjunction with our model to expand the basis of the model or to characterize a breast shape by condensing information down to a few principal components. There have been several other methods proposed to automatically segment or detect the breast boundaries in mammograms.[119-122]

Limitations of this study include the lack of including an automated nipple detection and removal algorithm, so that the model of the breast shape would not be

affected by its presence. In addition, inclusion of an automated pectoralis detection algorithm, especially in the MLO view, such as that proposed by Kowk et al.,[123] and its inclusion as another edge in the (x,y) coordinate pair vector for the PCA could allow for this image feature to also be included. Other future work would include the expansion of the PCA with more mammograms, allowing for the study of PCA parameter variability with patient characteristics (e.g. age, glandular density), and for the investigation of constraining their values to generate corresponding CC and MLO view mammograms.

4.5 Conclusion

The objective PCA models of breast shape presented here show great promise and could become useful tools for breast imaging applications, both for analyzing acquired images and for generating realistic breast shapes. Testing of the models' ability to replicate clinical breast shapes encountered in an independent database were successful, and the ease of generating random, clinically realistic breast shapes with the models was shown.

The ability to characterize a breast shape using just a few parameters is of great utility in many applications, but pertaining to this thesis, it is a major contribution towards bringing the X-ray scatter correction algorithm for DBT, described in chapter 2, to the clinical realm.

CHAPTER 5

PCA MODEL-BASED X-RAY SCATTER MAP LIBRARY

In this chapter I present a study on the use of a library of DBT X-ray scatter images, derived from the PCA model of breast shapes presented in chapter 4, with the scatter correction algorithm, described in chapter 2. I also investigate the effect this has on clinical DBT acquisitions that contain glandular masses and microcalcifications.

The PCA-based mammography breast shape model for the CC view was extended to DBT and used to generate compressed breast volumes that span most of the range of clinically encountered breast shapes and sizes. MC simulations of these breast volumes undergoing DBT were then performed to form a library of CC-view X-ray scatter maps. The application of the X-ray scatter algorithm using this library was evaluated in an observer study. While only one in four observers preferred the scatter corrected DBT images as a whole, all observers exhibited a preference for the SC images when the lesion examined was a glandular mass. When the lesion examined consisted of microcalcification clusters, the observers exhibited a preference for the uncorrected images.

5.1 Significance

As detailed in chapter 2, current tomosynthesis systems, including the Selenia Dimensions (Hologic Inc., Bedford, MA) used in this study, lack X-ray scatter reduction measures, be it in software or hardware. This leads to the inclusion of the entirety of the X-ray scatter signal in the tomosynthesis projections, unlike in mammography, where an anti-scatter grid oriented towards the fixed X-ray source lowers the ratio of the number of scattered X-rays incident on the detector to that of the number of incident primary (non-scattered) X-rays, or the scatter-to-primary ratio (SPR). Therefore, as has been previously

shown, in DBT the SPR can be as high as 1.6,[20] resulting in reconstruction artifacts and reduced contrast. [48-50] I have previously described a software-based X-ray scatter correction algorithm that is applied to the acquired tomosynthesis projections before reconstruction in chapter 2 that has shown great potential in improving image quality.

The study presented in this chapter leverages the fact that in DBT, the compressed breast is positioned in the same manner as in full field digital mammography (FFDM). In the case of the Selenia Dimensions system, the FFDM and DBT acquisition occur during the same breast compression. Thus, the principal component analysis (PCA) models of breast shapes undergoing mammography described in chapter 4 can be extended to models of 3D breast volumes undergoing DBT acquisition.

Here I present a study on the incorporation of PCA models of breast volumes into a software-based X-ray scatter correction algorithm, to reduce execution time. The incorporation of the PCA models allows for the possibility of using a pre-computed library of X-ray scatter maps, thereby eliminating the need to perform a case-specific and time-intensive Monte Carlo (MC) simulation with each new patient. The improved scatter correction algorithm was applied to 40 patient cases, and four radiologists were recruited for an observer study. The results of this study suggest that the incorporation of an X-ray scatter map library and the scatter correction algorithm overall exhibit the potential to improve the image quality of lesions that consist of masses but not of microcalcification clusters and to improve the clinical performance of DBT when its use is widespread.

5.2 Materials and Methods

The study presented here expands upon a previously developed software-based scatter reduction algorithm for DBT imaging by incorporating a pre-computed library of scatter maps. These maps, along with primary signal maps, are generated from MC simulations that utilize voxelized breast volumes created by an objective PCA model of breast shape. The MC simulations are implemented in C++, based on the Geant4 Monte

Carlo simulation toolkit,[54, 55] and are similar to those used by Feng et al[109] and Sechopoulos et al.[19-21] The scatter and primary images are combined into a scatter-to-primary ratio (SPR) map which is then registered to the acquired breast projections. The low-frequency scatter signal is then removed from the projections, which are subsequently noise filtered, as described in section 2.2 Lastly, the scatter corrected projections undergo filtered backprojection (FBP) 3D reconstruction.

This procedure was tested on 40 patient tomosynthesis acquisition sets, previously classified as BIRADS 4 or 5, that were acquired by the Selenia Dimensions DBT system. For the observer study, the radiologists were asked to evaluate the original and scatter corrected images for conspicuity of the lesions and detail of the features relevant to diagnosis.

5.2.1 Model of Breast Volume Undergoing Digital Breast Tomosynthesis

This study utilizes the objective PCA model of breast volumes undergoing cranio-caudal (CC) view FFDM described in chapter 4. As the image of the compressed breast shape undergoing mammography is identical to the central angle (0 degrees) projection of the compressed breast shape undergoing DBT, the mammography model forms the basis for the DBT model. The 3D breast volume (3DV) is built using the 2D breast shape (2DS) generated by the model from chapter 4 and an estimated curvature of the compressed breast volume. A script was written in IDL 8.0 (Exelis VIS, Boulder, CO) that generates a binary image stack that represents the modeled breast volume with a given compressed thickness (*Th*).

First, the 2DS is eroded to account for the magnification owing to the distance of half of *Th*, as the image of the 2DS is assumed to be that of the central slice of the voxelized breast volume. Building upon this central slice, two parabolic fits, one for the slices of the 3DV above and one for the slices below the central slice, are applied to represent the extent to which the 2DS is eroded. This erosion is necessary as the

curvature of the compressed breast is assumed such that the breast tissue extends furthest from the chest wall at the center, and is closer to the chest wall at the top and bottom (Figure 5.1). The parabolic fit below the center of the breast volume leads to slices in which the breast shape is eroded to a lesser degree than those above the center. This is a reasonable assumption to generate realistic breast volumes in which tissue distribution, as viewed laterally, is affected by gravity, which can be seen in Figure 5.1(b).

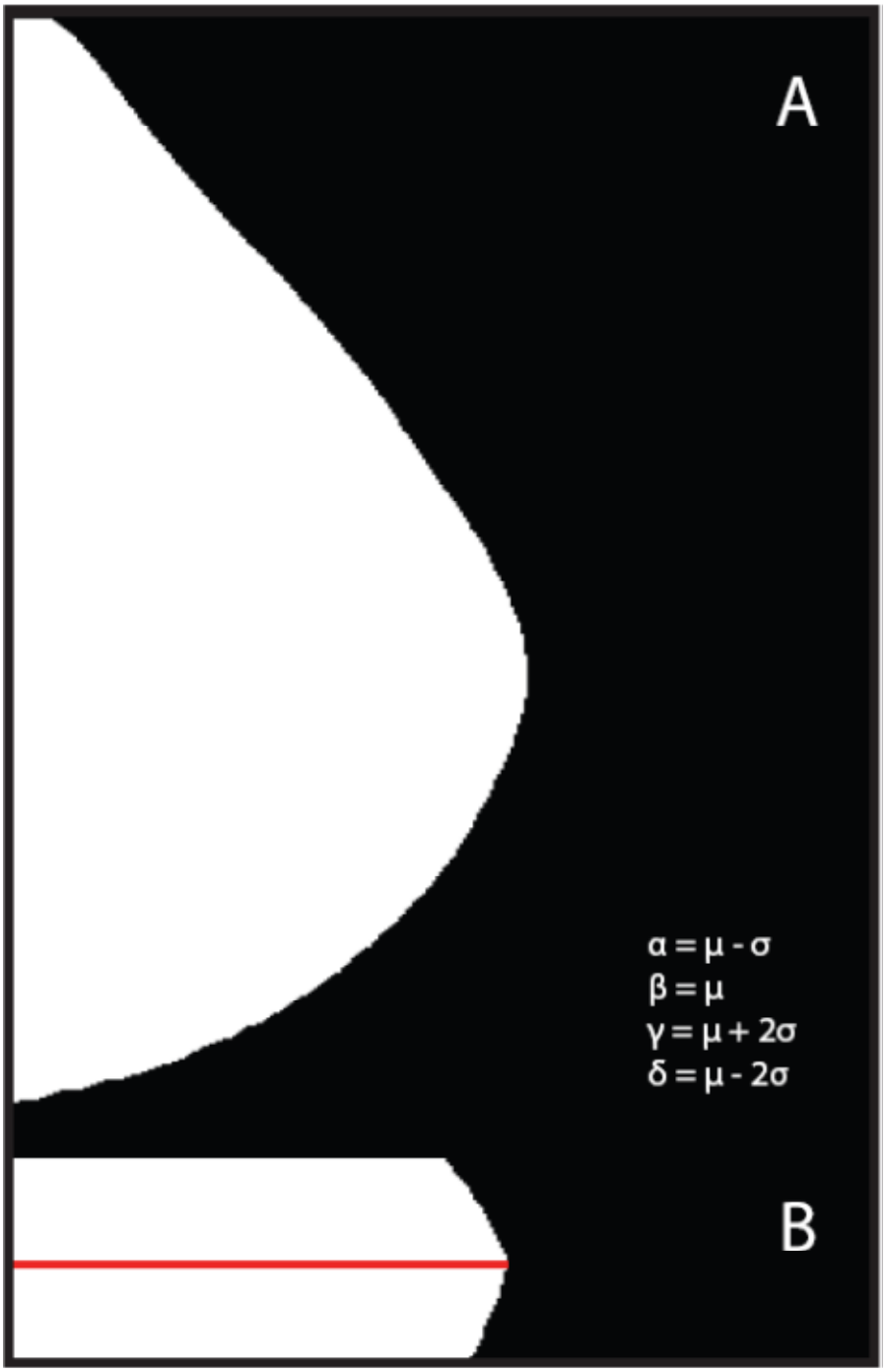

Figure 5.1 – Generation of the 3D Breast Volume from a 2D Breast Shape – (A) A top down view of the central slice of the volume generated from the PCA model breast shape of which the 4 principal components α, β, γ, and δ were each set to the indicated linear combinations of the component-specific mean (μ) and standard deviation (σ). (B) The side view of the 6cm breast volume from (A). Note that the curvature of the breast differs above and below the center slice. The red horizontal line indicates the slice location of the image shown in (A).

At the outset of this study, a 3DV library large enough to encompass the range of most clinically encountered breasts was desired. To create this library, the principal components of the 4-component PCA model were selected according to the values in Table 5.1. The mean (μ) and standard deviation (σ) of each principal component, as detailed in chapter 4, was used for these selections. As the first component, α, has been shown to be the greatest determinant of the breast shape and size, 5 values of α were chosen. For the other three components, only 3 values were chosen. *Th* was chosen to span the range from 2 cm to 8 cm, in 2 cm steps. Thus, a total of 540 distinct 3DVs were generated from the PCA model of breast volume undergoing DBT acquisition.

Table 5.2 – Selected Values for Principal Components of Library – To determine the input values to the PCA breast shape model, the mean (μ) and standard deviation (σ) of each component were used.

Table 5.1 Selected Values For Principal Components of Library

Principal Component	Number of Values	μ	σ	Selected Values
α	5	-3.185	15.78	μ + 2σ μ + σ μ μ - σ μ - 2σ
β	3	-0.296	8.225	μ + 2σ μ μ - 2σ
γ	3	-0.0654	3.658	μ + 2σ μ μ - 2σ
δ	3	-0.1286	1.994	μ + 2σ μ μ - 2σ

5.2.2 X-Ray Scatter Map Library

A voxelized version of the X-ray scatter MC simulation used by Feng et al[109] and described in chapter 2 was used to generate scatter image sets of all 540 3DVs from the PCA breast volume model undergoing DBT. The simulation is implemented in C++ and

matches the acquisition geometry of the Selenia Dimensions system for the CC view. The system acquires 15 projections over a 15 degree angular range, with the X-ray tube positioned directly perpendicular to the image detector for the central angle (tomosynthesis angle = 0 deg) projection. The detector measures 24 x 29 cm, with a pixel pitch of 70 μm. Complete details of the system have been previously published in Ren et al.[56] Thus, the MC simulated X-ray scatter map set of each 3DV consists of 15 scatter maps, one from each acquisition angle.

The glandular fraction of the breast for all simulations was set to 14.3%, which has previously been shown to be the population average.[75] This is a reasonable assumption as it has been shown that scatter varies to a minimal extent with glandular fraction.[20] Since it is known that X-ray scatter varies slowly over the detector, a pixel pitch of 0.7 mm was used in the MC simulations to decrease the simulation time by reducing the noise in the results.

5.2.3 X-Ray Scatter Correction

We obtained 40 patient cranio-caudal view DBT images classified as BIRADS 4 or 5. These images contained lesions, previously identified by a radiologist, evenly divided between masses and microcalcification clusters. The patient case datasets were acquired for an unrelated IRB-approved clinical study from which the images were released for use in other research projects.

The central angle projection of each acquisition set was subjected to the automated edge detection (section 4.2.2) and the PCA (section 4.2.3) previously described to obtain the PCA parameter vector $\boldsymbol{r}_{acq}$, which consists of the four principal components [α, β, γ, δ], of each image. From the scatter map library, the MC-simulated scatter map set of the 3DV with the closest matching PCA parameter vector was chosen.

To obtain noise-less X-ray primary images, raytracing[59] was performed to determine the pathlength of each ray from the source to each pixel on the image detector.

The pathlengths of the rays through each material in the simulation, combined with the attenuation coefficients for each material[60] and X-ray energy were used to calculate the estimated primary images. This allows for the calculation of a scatter-to-primary (SPR) map, which depicts the ratio of the scatter signal to the primary signal on a pixel by pixel basis, for each tomosynthesis acquisition projection. As a final step, the SPR map is resampled to the pixel pitch of the imaging system in DBT acquisition mode (140 μm).

The selected scatter map set and raytraced primary images were then used in the X-ray scatter correction algorithm described in chapter 2 to remove the scatter signal from the tomosynthesis acquisitions. The original (*O*) and scatter corrected (*SC*) images then underwent 3D reconstruction using the FBP reconstruction of the Selenia Dimensions system.

5.2.4 Observer Study

For this study, volumetric regions of interest (ROIs) centered on the lesion were isolated from *O* and *SC* reconstructions described above. Four radiologists experienced in breast imaging were recruited for this observer study. Each observer was presented with the *O* and *SC* DBT images side by side, randomized in both order and presentation, such as those depicted in Figure 5.2. The images are presented with comparable window width and level and observers could scroll through the volume stack. In addition, the overall brightness and contrast could be adjusted by the observer. Three observers completed the study in a single one hour session, while the other observer performed the study in two half hour sessions. The DBT reconstructions were displayed using the ImageJ image processing software (NIH, Bethesda, MD) on a DICOM calibrated[124, 125] 5 megapixel high-resolution flat-panel gray-scale dual-head display (DOME C5i, Planar Systems, Beaverton, OR).

Observers were asked to indicate their preference for either the left or right image on a scale from -5 (left) to +5 (right), with 0 indicating no preference, based on the

conspicuity of the lesion and the detail conferred by the image features. Ratings were normalized such that a preference for the *O* image is negative, whilst a preference for the *SC* image is positive. These ratings were then averaged to determine if each clinician exhibited a preference for either the *O* or *SC* images.

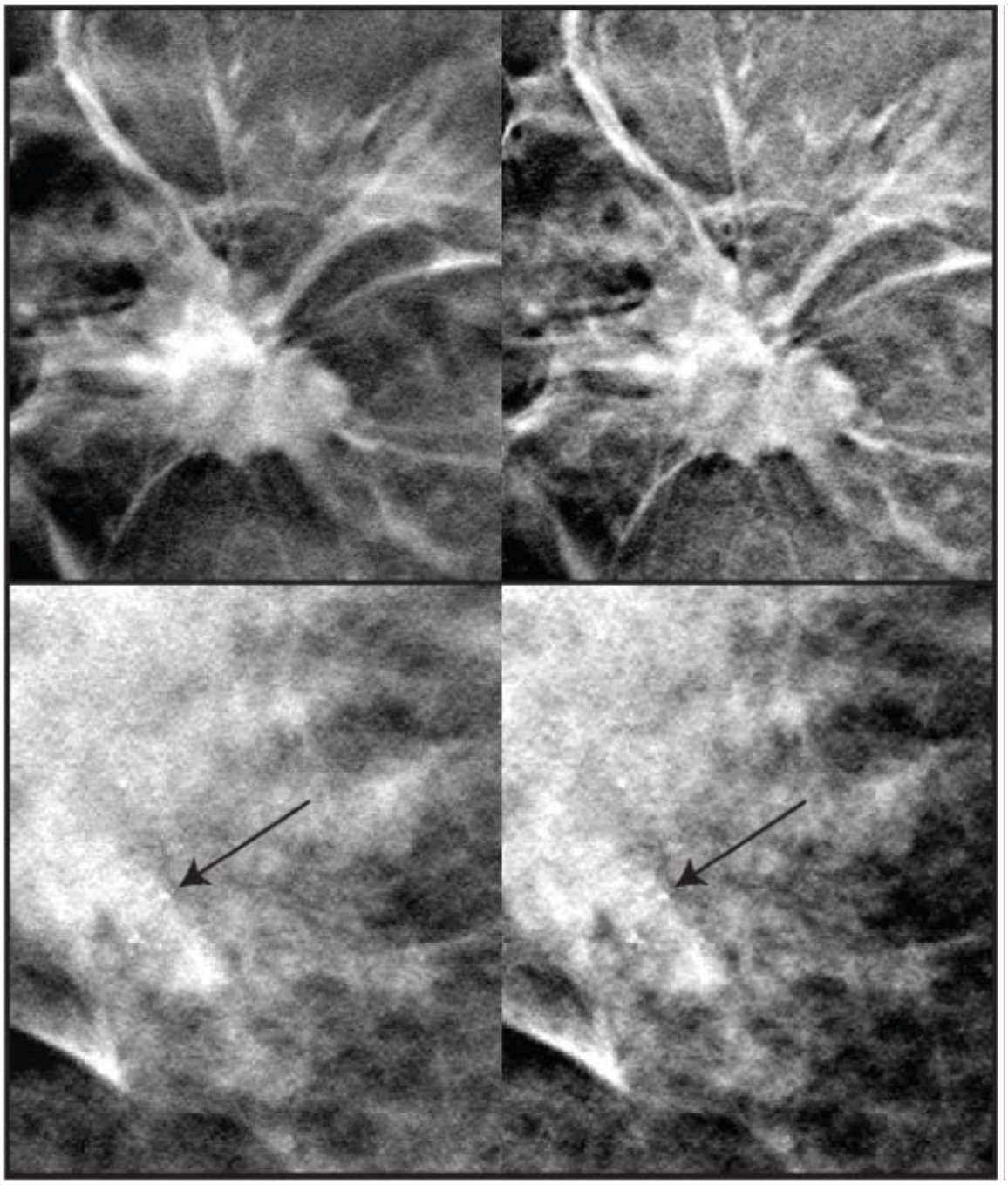

Figure 5.2 – Scatter Correction Observer Study – Shown here are two distinct DBT patient images. Observers were presented with these images, which consist of regions of interest centered on lesions, such as a glandular mass (top) and microcalcification clusters (bottom, arrow).

5.3 Results

5.3.1 X-Ray Scatter Map Library

The addition of the breast volume PCA model and the X-ray scatter map library to the scatter correction algorithm successfully reduced the execution time of the algorithm from hours to minutes. After all 40 patient cases were subjected to the automated edge detection and PCA, it was discovered that the initial library of 3DVs generated in section 5.2.1 was insufficient in terms of range and granularity. For *Th*, 39 of the 40 patient images spanned the range of 2 cm to 8 cm, but it was determined that a smaller step size of 1 cm is required. One patient case had a *Th* of 9 cm. Additional 3DVs were generated to match those patient cases and additional MC simulations were performed. It was determined that a library would need to consist of scatter map sets of 3DVs characterized by the principal component values listed in Table 5.2. As a total of 16,087 scatter maps sets would be needed to complete this library for the CC view for this specific DBT system, an estimated 210,000 hours of computing time and 120 gigabytes of storage space would be required.

Table 5.2 – Range and Granularity of Principal Components of Library – Shown here are the values for the PCA breast shape model required for a fully developed X-ray scatter map library. The mean (μ) and standard deviation (σ) of each component were used to determine the values.

Table 5.2 Range and Granularity of Principal Components of Library

Principal Component	Number of Values	μ	σ	Selected Values
α	7	-3.1852	15.7792	μ + 4σ μ + 3σ μ + 2σ μ + σ μ μ - σ μ - 2σ
β	7	-0.2959	8.2254	μ + 4σ μ + 3σ μ + 2σ μ + σ μ μ - σ μ - 2σ
γ	7	-0.0654	3.6579	μ + 4σ μ + 3σ μ + 2σ μ + σ μ μ - σ μ - 2σ
δ	7	-0.1286	1.9939	μ + 3σ μ + 2σ μ + σ μ μ - σ μ - 2σ μ - 3σ

5.3.2 Observer Study

The combined total mean score for all images and observers is 0.169±0.37 ($p > 0.39$), which indicates that overall, there is no preference for either the *O* or *SC* images (Table 5.3). The combined mean score given for images that contained a soft tissue mass is 1.06±0.45 ($p < 0.0001$), with every observer averaging positive scores for these images. However, when the lesion in the image consists of microcalcification clusters,

the combined mean score given to these images is -0.725±0.51 ($p < 0.009$), which indicates that for these lesions the observers prefer the *O* images.

Also of note is that one radiologist, Observer A, consistently prefers the *SC* images, with a total mean score of 1.575±0.45 ($p < 0.002$), and positive mean scores for masses and microcalcifications.

Table 5.3 – Observer Study Results – Shown here are the mean scores for each observer for the images with a glandular mass, microcalcification clusters, and the combined case total, as well as the total for all observers. Positive scores indicate a preference for the scatter corrected images, while negative scores indicate a preference for the uncorrected images.

Table 5.3 Observer Study Results

Observer	Mean Total Score	Mean Mass Score	Mean Microcalcification Score
A	1.575	2.6	0.55
B	-0.175	0.05	-0.4
C	-0.375	0.55	-1.3
D	-0.35	1.05	-1.75
Combined	0.169	1.06	-0.725

5.4 Discussion

The study presented in this chapter shows that a library of X-ray scatter maps computed from MC simulations that incorporate compressed breast volumes generated by a PCA model can be successfully integrated into a DBT scatter correction algorithm, which can reduce the processing time by two orders of magnitude. In addition, the incorporation of the PCA model from chapter 4 into the algorithm, and the testing on 40 DBT patient cases, led to the delineation of the parameter space that such a library must cover (Table 5.2). Thus, a total of 16,087 scatter map sets would be needed to complete this library for CC view DBT.

The required expansion of the initial scatter map library from section 5.2.2 suggests that the PCA model of breast shape detailed in chapter 4 may need to be

expanded or revised. Firstly, the mammographic images that formed the basis of that model measure 19.14 x 22.94 cm, while the detector of the Selenia Dimensions measures 24 x 29 cm. It is possible that the breadth of the FFDM foundation database of the PCA model is limited by the detector size of the Senographe 2000D (GE Healthcare, UK) mammography system. These cases have a mean *Th*=5.91 cm and the average calculated value of the first principal component, α, which was shown to be the greatest determination of 2DS size, is more than one standard deviation greater than the mean (average(α) = μ + 1.04σ) from the PCA model. An expansion of the set of foundation images of the PCA model could be performed in the future by including mammograms and central angle DBT projections acquired by the Selenia Dimensions system.

The observer study with the *O* and *SC* reconstructions reveals that the scatter correction algorithm has the potential to improve the conspicuity and image quality of masses in DBT images. The contrast enhancement of masses conferred by the algorithm was also previously shown in phantoms by Feng et al,[109] and is detailed in chapter 2. The observers' preference for the *O* reconstructions when presented with images depicting microcalcification clusters could be due to the fact that the paired images were displayed with comparable windows in small ROIs. In addition, microcalcifications already exhibit high contrast in the *O* reconstructions, and could be also obscured by contrast-enhanced glandular tissue in the *SC* reconstructions. It is important to note that this study is limited to conspicuity of lesions, small ROIs, and to BIRADS 4 and 5 images. A larger observer detection study that includes negative patient cases, and in which the observers are asked to evaluate the reconstructions as a whole, would need to be performed to better evaluate the clinical impact of the algorithm.

And finally, this study focused on CC view DBT images. A similar study would need to be performed to accommodate the medio-lateral oblique (MLO) view. A PCA model of breast shapes undergoing MLO mammographic imaging is already complete, and is detailed in chapter 4.

5.5 Conclusion

The incorporation of the CC view PCA model of breast shape into the scatter correction algorithm has improved the efficiency of the algorithm and provided insight into the size and range required of a pre-computed CC view X-ray scatter map library. The observer study presented in this chapter is also the first test of the scatter correction algorithm with patient images and human observers, and demonstrates the potential it has to improve the clinical performance of DBT.

CHAPTER 6

SPECTRAL RECONSTRUCTION ALGORITHM

This chapter describes the optimization and testing of a novel spectral reconstruction method and the feasibility of using this method with a new tube voltage-switching DBT acquisition technique. The spectral reconstruction (SR) algorithm incorporates the varying attenuation properties of the full spectrum of X-ray energies generated by the X-ray tube and used to acquire the tomosynthesis projections.

The SR algorithm was tested by using it to reconstruct images of a breast tomosynthesis phantom and patient images. The image quality of the SR phantom and patient images was higher than that demonstrated by MLEM reconstructions. The tissue patterns of the heterogeneous tomosynthesis phantom were more clearly defined and beam hardening artifacts were greatly reduced. The images reconstructed with the SR algorithm also demonstrated an increase in SDNR. This study demonstrates that the SR algorithm has the potential to greatly enhance the image quality of breast tomosynthesis images and reduce reconstruction artifacts. The algorithm also offers the potential to allow for new image acquisition techniques, such as DBT acquisition with varying X-ray beams, which could allow for a reduction in patient dose. This may result in an increase in early breast cancer detection and a reduction in unnecessary additional procedures to investigate false positives.

6.1 Significance

Conventional DBT reconstruction algorithms are based on an underlying assumption that the energy of every X-ray that are incident on the detector is identical. This includes the most commonly used reconstruction algorithms: filtered backprojection (FBP), or iterative methods such as simultaneous iterative reconstruction technique

(SIRT) and maximum likelihood expectation maximization method (MLEM).[15, 126] This is also true of novel reconstruction algorithms that have been proposed for DBT with varying levels of testing, success, and development, including regularization methods such as total variation regularization,[127, 128] total *p* variation regularization,[129-131] joint entropy regularization,[132, 133] and adaptive diffusion regularization.[134] Even iterative methods such as iterative penalized maximum likelihood,[135] iterative maximum *a posteriori* statistical reconstruction,[136] and Bayesian inference[137] are also based on this same underlying assumption.

While this mono-energetic assumption is adequate for imaging applications which utilize a high energy X-ray beam with heavy filtering, such as computed tomography (CT), the beams typically used in DBT imaging are comprised of X-rays of a wider and lower range of energies. Beam hardening is thus a much greater concern, as the distribution of X-ray energies that are incident on the image detector differs greatly from that of the beam entering the breast. The only way to account for this phenomenon is to reconstruct the DBT 3D images using a revolutionary spectral reconstruction (SR) algorithm, such as the one developed by a collaboration of our group and Dr. James Nagy's group that takes into account the distribution of X-ray energies present in the beam and the different attenuation that X-rays of different energies undergo.[51] This algorithm was then expanded upon by Bustamante et al, by generalizing the mathematical model for material decomposition of the reconstruction volume.[52]

In this study, a GPU implementation of this algorithm, is optimized and tested with DBT patient images. In addition, the spectral reconstruction algorithm was modified to reconstruct images from a dual-spectrum single-acquisition of DBT projections; that is a set of projections acquired using two different X-ray spectra. This has the potential to improve image quality or reduce radiation dose without compromising image quality. It has previously been shown that the use of a second, higher energy X-ray spectrum to

acquire half of the DBT projections would result in a lower total radiation dose to the imaged breast,[19] which is a very important consideration.

6.2 Materials and Methods

The SR algorithm described here incorporates a spectral model for the X-ray source and a statistical model for the reconstructed volume. The projected images are then estimated using this model of the volume and by modeling the nonlinear energy transmission radiography process. A gradient descent search algorithm is then used to calculate the best solution for the reconstruction volume. A GPU implementation of this algorithm, developed in C++ with OpenCL libraries, was first optimized for reconstruction of images acquired by the Selenia Dimensions (Hologic Inc., Bedford, MA) DBT system. The SR algorithm was then used to reconstruct images of a heterogeneous breast tomosynthesis phantom, and the image quality of the reconstructions was evaluated. Subsequently, a simulated dual-spectrum single-acquisition imaging method was tested on a phantom. The dual-spectrum images were reconstructed using the SR algorithm and the image quality of the reconstructions was also evaluated. Finally, leveraging the results of the dosimetric characterization of the Selenia Dimensions detailed in chapter 3, the resulting radiation dose of the dual-spectrum single-acquisition imaging method was calculated.

6.2.1 Spectral Reconstruction Algorithm

In DBT imaging and computed tomography, the imaged 3D volume, or breast tissue in this case, may be discretized into a regular grid of *N* voxels while the acquired projections are discretized into a regular grid of *M* pixels. The voxels are described by a multimaterial model which incorporating the linear attenuation coefficient and weight fraction of each material. Under the assumption that the densities of all materials are similar, the linear attenuation coefficient $\mu(e,j)$ for each X-ray energy *e* of a single voxel *j* composed of *D* distinct materials can be approximated as:

$$\mu(e,j) \approx \sum_{m=1}^{D} w_m(j) * c_m(e) \qquad \text{(Equation 6.1)}$$

where $w_m(j)$ represents the weight fraction in voxel *j* of the *m*th material, and $c_m(e)$ are the linear attenuation coefficients for the *m*th material at X-ray energy *e*. Then for each acquisition projection angle, we compute the sum of the monochromatic X-ray traces over all energies to obtain an estimate of the projection b for acquisition angle θ:

$$\mathbf{b}_\theta = \eta_\theta + \sum_{e=1}^{E} \Psi(e) * \exp\left(-\sum_{j=1}^{N} a(i,j) \sum^{D} w_m(j) * c_m(e) \right) \qquad \text{(Equation 6.2)}$$

where $a(i,j)$ represents the length of the ray that passes through voxel *j* and is incident upon pixel *i*, $\Psi(e)$ is the energy fluence, and η_θ represents additive noise and X-ray scatter. A gradient descent search algorithm is then used to minimize the error between the estimate **b** and the acquired projection by updating the weight fractions $w_m(j)$ of all voxels in the volume. Full details of the algorithm can be found in Chung et al and Bustamante et al.[51, 52]

6.2.2 Spectral Reconstruction of Phantom Images

DBT image projections of a 5 cm thick heterogeneous Model 020 BR3D breast tomosynthesis phantom (CIRS Inc., Norfolk, VA) were acquired using the Selenia Dimensions, and reconstructed using the SR algorithm and using a mono-energetic

maximum likelihood expectation maximization (MLEM) reconstruction algorithm.[53] The phantom consists of 1 cm thick slabs of a heterogeneous 50%/50% mixture of adipose and glandular tissue. This phantom represents real tissue better than a homogenous phantom, and tests tomosynthesis' ability to suppress tissue superposition. The 1cm thick slab located at the center of the phantom, the "target slab", has embedded spheroid tumor masses, fibers, and microcalcification clusters of differing sizes (Figure 6.1).

In addition, two distinct patient cases were also reconstructed using both methods. The patient case datasets were acquired for an unrelated IRB-approved clinical study from which the images were released for use in other research projects. The datasets are all reconstructed to image stacks of 0.14 mm x 0.14 mm x 1.0 mm voxels.

Image quality was assessed by comparing the SR and MLEM reconstructions, in terms of the swirl patterns of the two distinct materials in the phantom, and the presence of beam hardening artifacts caused by microcalcification clusters. Quantitative analysis was performed by measuring and comparing the signal difference to noise ratio (SDNR) of three spherical glandular masses embedded in the target plate of the phantom.

All reconstructions using the SR algorithm were performed on a Tesla C2070 GPU (NVidia Corp., Santa Clara, CA), while the MLEM reconstructions, implemented as a distributed computing task in IDL 8.0 (Exelis VIS, Boulder, CO), were performed on a 22-node computing cluster.

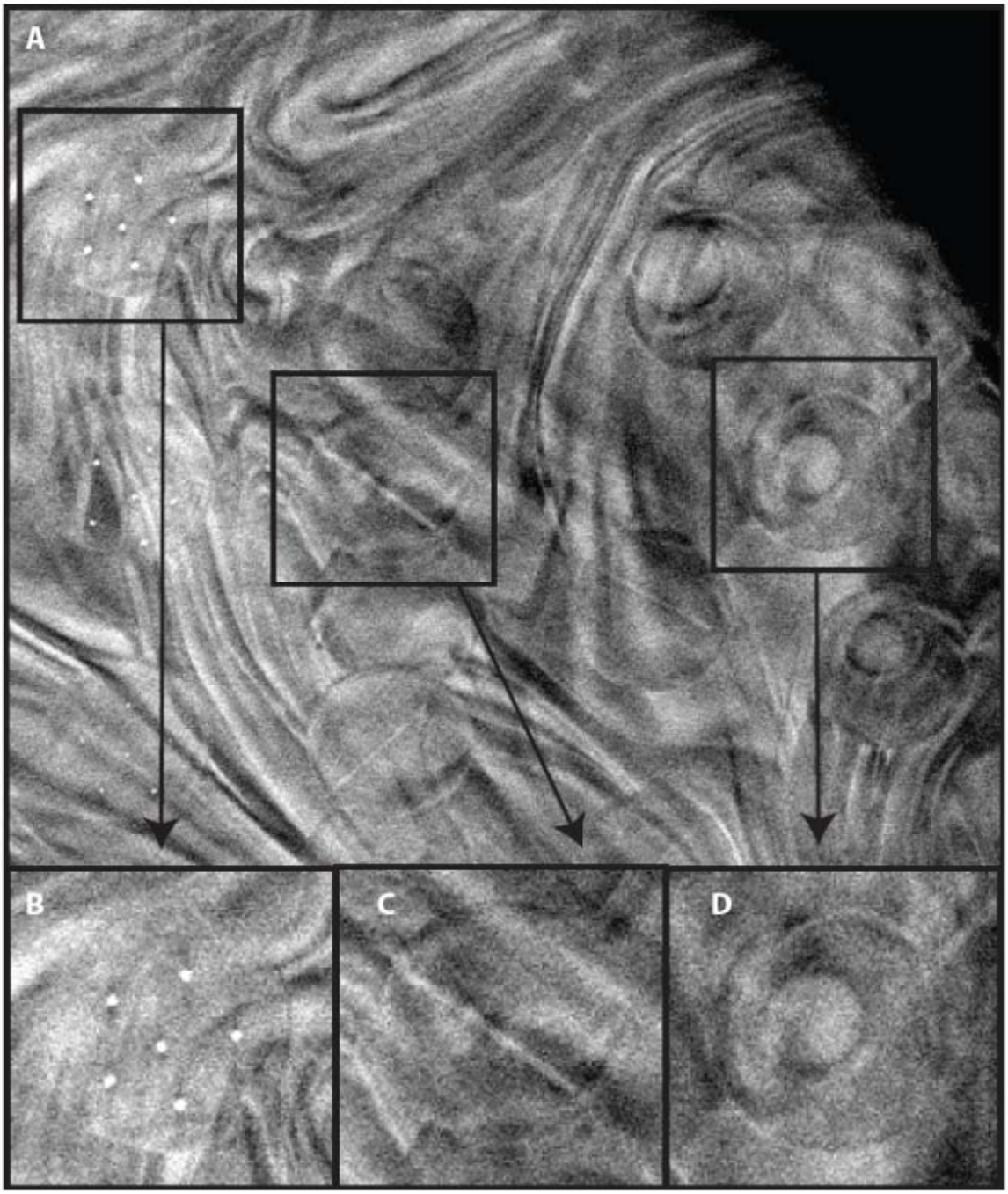

Figure 6.1 – (a) The CIRS Model 082 Breast Mammography Phantom, representing a 50/50 heterogeneous mixture of glandular and adipose tissue contains three types of embedded targets of varying sizes: (b) spheroid masses, (c) fibers, and (d) microcalcification clusters.

6.2.3 Simulated Dual-Spectrum Acquisition

As DBT systems are currently unable to perform X-ray tube voltage switching, two separate DBT acquisitions, with distinct X-ray energy spectra, were performed of another 5 cm thick heterogeneous breast tomosynthesis phantom. The phantom consists of a custom 1 cm thick 100% adipose plate in the center with slots for customizable targets. For this study, targets of differing glandular fractions and targets containing microcalcification clusters were used. Two 1cm thick slabs of the CIRS Model 082 phantom are placed above and below the target plate to complete the phantom. This phantom was then imaged once with the Selenia Dimension's AEC designed X-ray tube voltage and tube current-exposure time product. These projections form the Base set. The second acquisition was performed with a manual selection of a higher tube voltage with added filtration (.254 mm Cu), and a tube current-exposure time product that produced a similar signal at an ROI located at the center of the phantom in the central angle DBT projection. These projections form the High Energy set.

The two X-ray spectra used for the Base and High Energy acquisitions were modeled according to the method described by Boone et al,[57] and detailed in section 3.2.2 of this thesis. These spectra models are inputs into the SR algorithm. The exposure, in mR, output by the X-ray tube for each spectrum was measured using a calibrated ACCU-DOSE dosimeter (Radcal Corp, Monrovia, CA) and a dedicated mammography ionization chamber (Model 10X6-6M, Radcal Corp, Monrovia, CA).

The simulated dual-spectrum single-acquisition (DS) set of projections was then formed by combining alternating projections from the Base set and the High Energy set. Eight projections were selected from the Base set, including the central angle projection, and seven were selected from the High Energy set. The Base and DS sets of projections underwent reconstruction using the SR algorithm, while the Base set also underwent

filtered back projection (FBP) reconstruction, as this is the default method of the Selenia Dimensions system. The SDNR of targets of differing glandular fractions was assessed.

6.2.4 Dual Spectrum Dosimetry

The resulting radiation dose of the simulated dual-spectrum acquisition and of the Base acquisition was calculated following the method detailed in chapter 3. Monte Carlo (MC) simulations were performed to determine the normalized mean glandular dose, on a per projection basis, in units of milliGray per milliGray air kerma for both the Base acquisition (D_gN_B) and the High Energy acquisition (D_gN_{HE}). For the Base acquisition, the entrance exposure measured by the dosimeter was used in conjunction with the D_gN_B to calculate the total mean glandular dose (MGD_B). For the total mean glandular dose of the DS acquisition (MGD_{DS}), the total mean glandular dose was calculated on a per projection basis using both D_gN_B and D_gN_{HE}.

6.3 Results

6.3.1 Spectral Reconstruction of Phantom Images

The image quality of the reconstructions generated by the SR algorithm is higher than that exhibited by images generated by the MLEM reconstruction algorithm. This can be seen in both the phantom images and the patient images. In the case of the phantom images, visual inspection reveals that the heterogeneous glandular/adipose tissue swirl patterns are more clearly defined in the SR images (Figure 6.2).

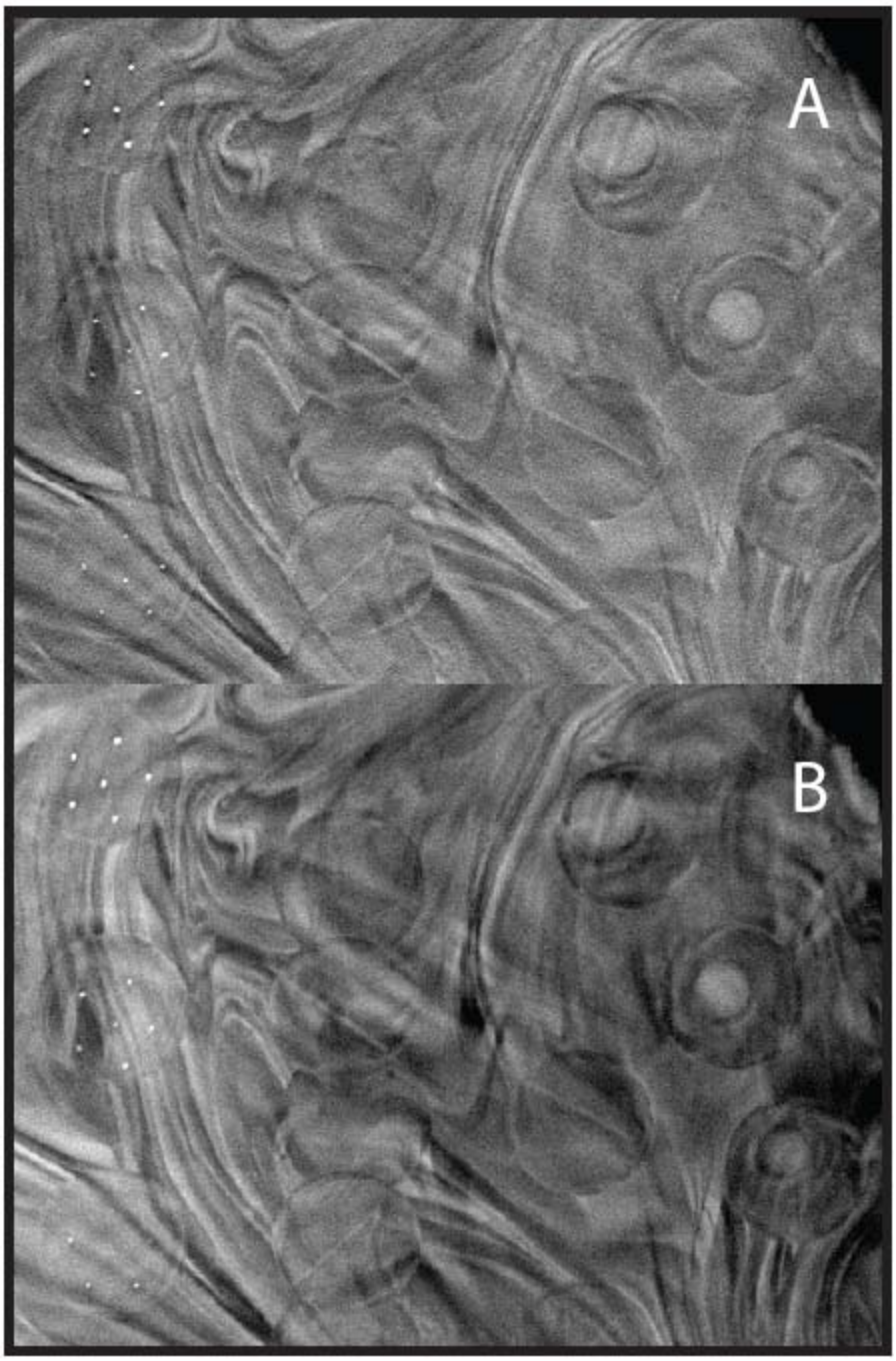

Figure 6.2 – Reconstructions of Phantom Images – (A) Center slice of the MLEM reconstruction of the phantom images. (B) Center slice of the SR reconstruction of the phantom images. Microcalcification clusters (left), fibers (middle), and glandular masses (right) can be seem in the phantom. These images are shown with comparable window widths but different levels.

Quantitative analysis of the embedded spheroid masses in the reconstructions reveals that SDNR of the masses is higher in the SR images than in the MLEM images (Table 6.1). The average improvement in SDNR is 71% ($p = 0.02$). The smallest embedded mass visible, C, in particular, is much easier to distinguish in the SR images (Figure 6.3A). The fibers are also easier to identify in the SR images (Figure 6.3B). And the beam hardening artifacts seen in the MLEM reconstructions are greatly reduced in the SR images (Figure 6.3C).

Table 6.1 – Signal Difference to Noise Ratio of Glandular Masses – Comparison of the image quality of the spherical embedded masses in the phantom between the spectral reconstruction (SR) images and the maximum likelihood estimation maximization (MLEM) images.

Table 6.1 Signal Difference to Noise Ratio of Glandular Masses

	SDNR		
Embedded Glandular Mass	MLEM	SR	SR/MLEM
A	1.83	3.33	1.82
B	1.33	2.33	1.75
C	3.00	4.67	1.56

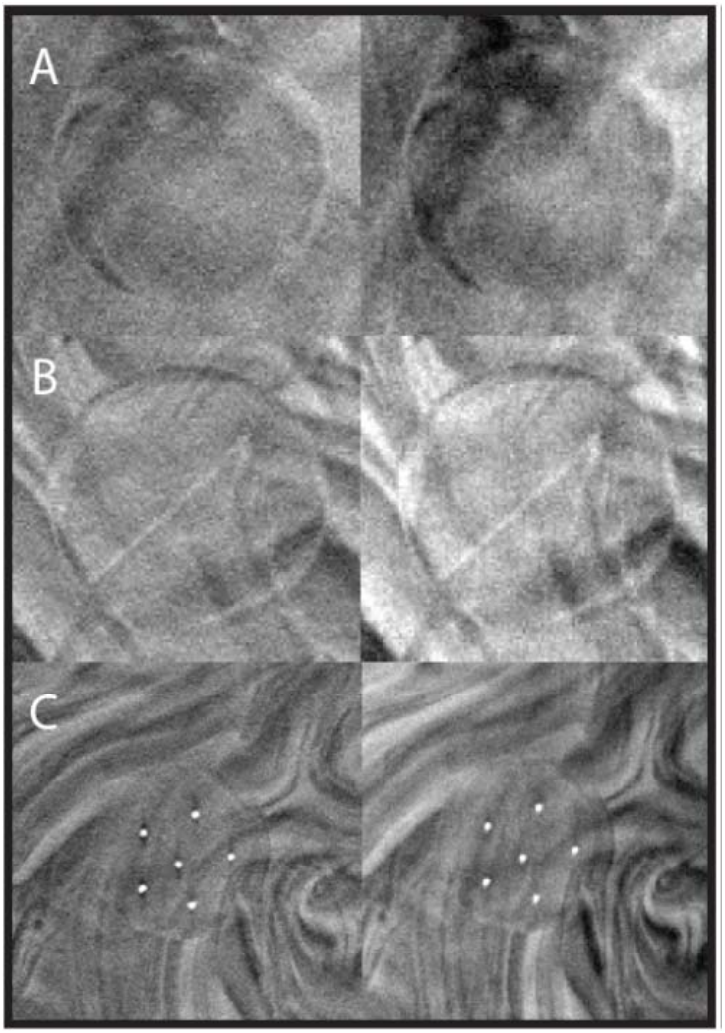

Figure 6.3 – ROIs of Reconstructions of Phantom Images – Zoomed in views of the smallest embedded glandular mass D (A), a fiber (B), and a group of microcalcification clusters (C) of the central slices of the MLEM (left) and SR (right) reconstructions. These images are shown with comparable window widths but different levels.

The image quality of the reconstructed patient images generated by the SR algorithm is also higher than that exhibited by images generated by the MLEM reconstruction algorithm. Visual inspection revealed that the microcalcifications were easier to distinguish in the SR images, and that the beam hardening artifacts in the direction of the X-ray tube motion are diminished in the SR images (Figure 6.4).

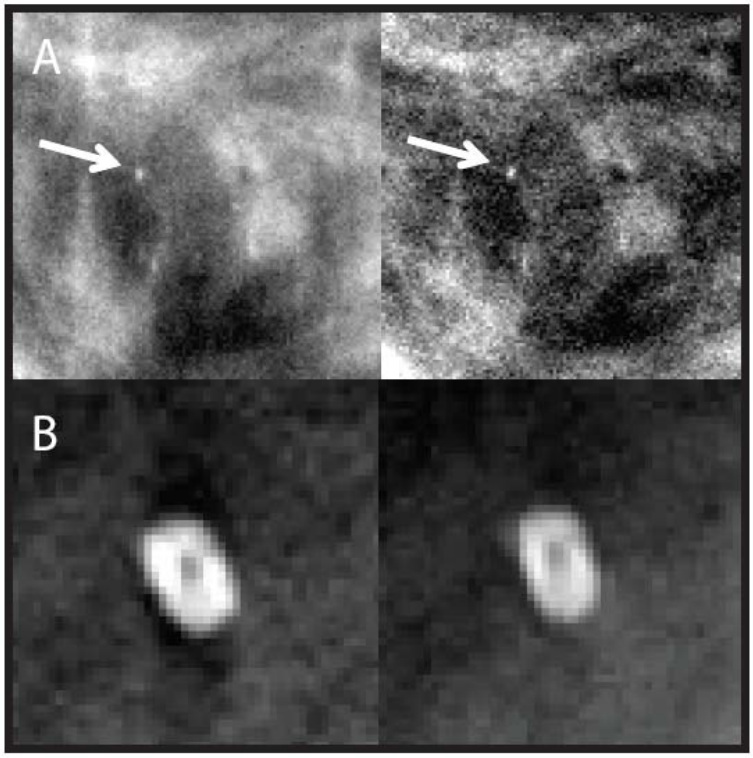

Figure 6.4 – ROIs of Reconstructions of Patient Images – Shown here are regions of interest of two patient images in which microcalcification clusters are present. In the first case (A), image contrast in higher, and the microcalcification (white arrow) is easier to distinguish in the SR images (right) than in the MLEM reconstructions (left). In the second case (B), the beam hardening artifact produced by the mono-energetic MLEM reconstruction (left) is greatly reduced in the SR image (right). These images are shown with comparable window widths but different levels.

6.3.2 Spectral Reconstruction of Dual-Spectrum Images

The two X-ray spectra used for the Base and High Energy acquisitions are described in Table 6.2, and their models are shown in Figure 6.5.

Table 6.2 – X-ray Tube Voltage and Tube Current-Exposure Time Products – These are the settings chosen for the DBT acquisitions in the dual-spectrum study

Table 6.2 X-ray Tube Voltage and Tube Current-Exposure Time Products

Acquisition	Tube Voltage (kVp)	Total Tube Current-Exposure Time Product (mAs)	Added Filtration	Entrance Exposure (mR)
Base	31	54	None	544
High Energy	49	30	Cu (0.254 mm)	37.8

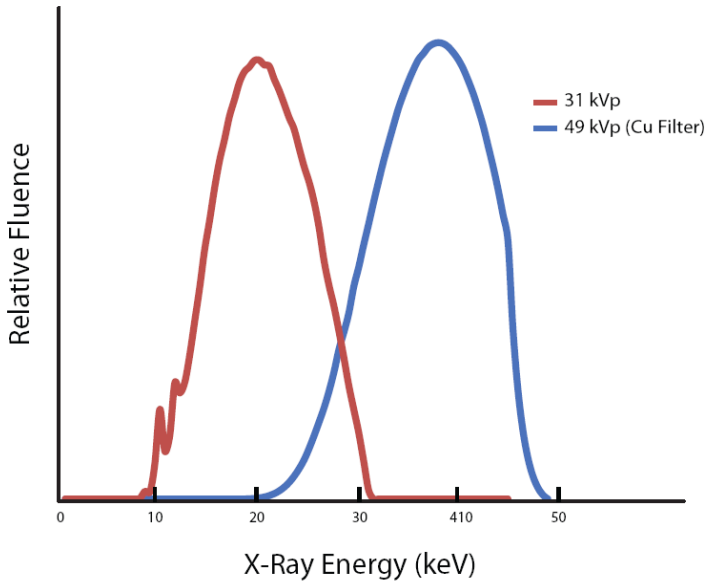

Figure 6.5 – Modeled X-ray Spectra - Depicted here is the relative fluence of the two X-ray spectra described in Table 6.1, the 31 kVp (red) and the 49 kVP (blue) with added Cu filtration.

The image quality of the DS reconstructions is slightly lower than that observed in the SR reconstructions, which in turn is comparable to that of the system's FBP reconstructions. The SDNR of the targets of multiple glandular fractions is shown in Table 6.3. The average change in SDNR between the FBP and SR images is -8.18 ± 11.6% ($p>0.52$), while the average change in SDNR between the SR and DS images is -16.0 ± 9.25% ($p>0.08$). Center slices of the FBP reconstruction, the SR reconstruction of

the Base set of projections, and the SR reconstruction of the DS set are shown in Figure 6.6. Zoomed in views of a group of microcalcification clusters show that the beam-hardening artifacts visible in the FBP images are absent from the SR Base images and SR DS images (Figure 6.7).

Table 6.3 – Signal Difference to Noise Ratio of Dual Spectrum Images – Comparison of the image quality of the targets of differing glandular fractions in the phantom between the filtered backprojection (FBP) reconstructions, and the spectral reconstructions of the Base set of projections and the simulated dual-spectrum (DS) set.

Table 6.3 Signal Difference to Noise Ratio of Dual Spectrum Images

	SDNR		
Target Glandular Fraction (%)	FBP	Base	DS
0	0.31	0.33	0.44
20	1.31	1.29	1.15
40	2.61	1.38	1.72
60	2.98	2.56	2.16
80	3.32	2.38	1.67
100	3.35	4.50	3.46

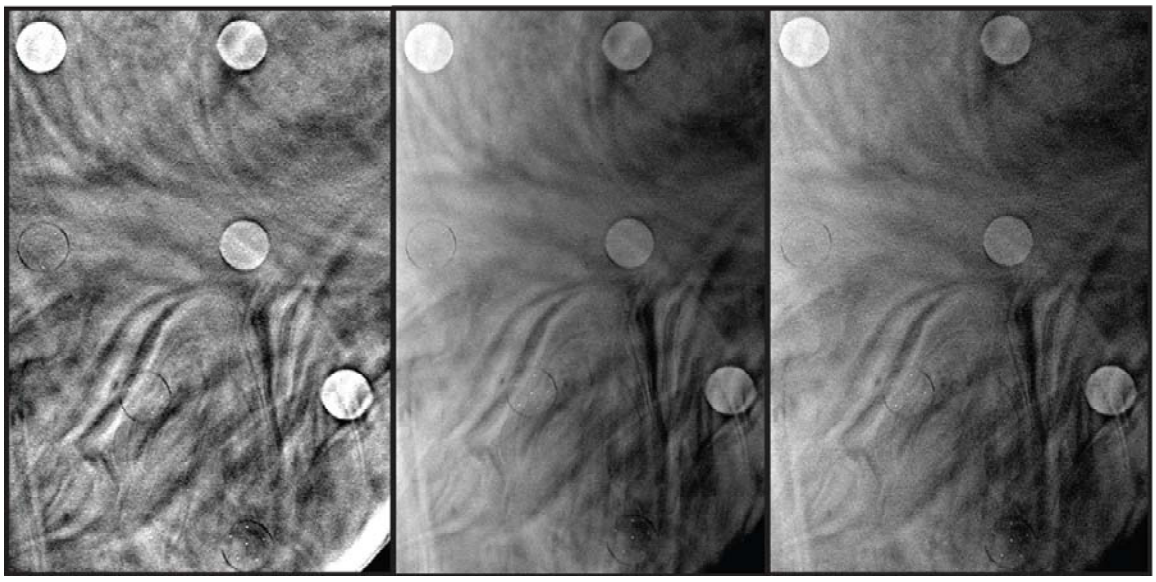

Figure 6.6 – DBT Reconstructions of the Simulated Dual Spectrum Acquisition – Center slices of the FBP reconstructions (left) and the SR reconstructions of the Base set of projections(middle), and the SR reconstructions of the DS set (right) are shown here. The image quality of the DS reconstructions is comparable to that of the Base set. These images are shown with comparable window widths but different levels.

Figure 6.7 – Zoomed-In Views of the DBT Reconstructions – These are zoomed in views of the images shown in Figure 6.6, centered on a group of microcalcification clusters. Beam hardening artifacts around the microcalcifications can be seen in the FBP reconstructions (left). However, neither the SR reconstructions of the Base set of projections (middle), nor the SR reconstructions of the DS set (right) exhibit such artifacts. The image quality of the DS reconstructions is comparable to that of the Base set. These images are shown with comparable window widths but different levels.

6.3.3 Dual-Spectrum Dosimetry

The calculated MGD_B is 1.12 mGy, while the MGD_{DS} would be 0.73 mGy, a 35% reduction in dose (Table 6.4).

Table 6.4 – Radiation Dose Resulting from Dual-Spectrum Imaging – the normalized mean glandular dose (DgN) and breast skin entrance exposure are used to calculate the total mean glandular dose (MGD) of each type of acquisition

Table 6.4 Radiation Dose Resulting from Dual-Spectrum Imaging

Acquisition	Tube Voltage (kVp)	DgN (mGy/mGy air kerma)	Entrance Exposure (mR)	MGD (mGy)
Base	31	54	544	1.12
High Energy	49	30	37.8	0.325
Dual Spectrum	31/49	-	308	0.73

6.4 Discussion

The results of this study suggest that the SR algorithm has the potential to offer improved DBT image quality over a mono-energetic reconstruction such as MLEM, as illustrated by the statistically significant improvements in SDNR of the embedded spherical masses in the phantom images. Visual inspection of the phantom images and the patient images also supports this conclusion. The heterogeneous swirl patterns of the phantom appear more distinguished in the SR images in Figure 6.2. The SR images of microcalcification clusters in the phantom also exhibit reduced beam-hardening artifacts, as can be seen in Figure 6.3C. The contrast of the smallest glandular mass visible, as seen in Figure 6.3A also appears to be enhanced in the SR images.

The patient images also appear to exhibit improved contrast in the SR images, as displayed in Figure 6.4. Also evident in those images is the reduction of beam-hardening artifacts around microcalcifications. This artifact reduction is achievable due to the poly-energetic nature of this method. Due to the small number of patient images presented in

this study, future work would need to include an observer study to compare a larger number of SR and MLEM images. It is also not yet clear how the SR algorithm would impact clinical performance. In addition, the GPU implementation of the SR algorithm is very efficient, as all of the images presented in this chapter were reconstructed within minutes, which is important in the clinical realm.

This study also reveals that dual-spectrum single-acquisition DBT imaging is feasible, as such a technique could reduce patient dose by a nontrivial amount without compromising image quality. This dose reduction is realized because the projections acquired with the higher energy spectrum result in a much lower entrance skin exposure to achieve the same signal at the detector. However, it appears that the SR algorithm in this case offers lower image quality than the FBP reconstruction of the Selenia Dimensions system. This could be explained by the proprietary post-processing performed by the machine that is unavailable to the public. In addition, the lower quantum efficiency of the detector at higher energies could be contributing to the lower SDNR for the higher energy projections.

It is important to note that this dual-spectrum feasibility study is also limited to just two different X-ray spectra. Future work would need to be performed to optimize the different spectra to be used, in terms of both tube voltage and added filtration. The subsequent differences in patient dose would also need to be evaluated when optimizing the spectra. Following that, an observer study or a clinical imaging trial would need to be performed to determine the impact on detection and diagnosis of breast cancer of this new acquisition and reconstruction method. It should be noted that if this dual-spectra acquisition method is shown to be successful, for it to be translated to clinical use, DBT systems would have to include the capability to execute fast, automated tube voltage switching during the projection acquisition process, but technology to this effect can be easily implemented by the tomosynthesis system manufacturers, so this acquisition method is clinically relevant.

In addition, further follow-up work would be to investigate the possibility of harnessing the ability of the SR algorithm to handle projections acquired with two different X-ray spectra to perform contrast enhanced breast tomosynthesis (CEBT) with the acquisition of a single set of DBT projections. Currently, CEBT is proposed to be performed either by temporal subtraction (acquiring one DBT projection set before and one after injection of the contrast agent and then subtracting the resulting reconstructions) or by dual energy (acquiring both DBT projection sets after injection, but one with a low energy X-ray spectrum and another with a high energy X-ray spectrum and then subtracting the reconstructions).[138, 139] In either case, two complete DBT acquisitions are performed, with the resulting increase in radiation dose and, in the case of temporal subtraction, with tissue mis-registration issues. With the adequate modification of our SR algorithm, it may be possible to obtain an iodine-only image reconstruction from a single set of projections, again intermittently varying the X-ray spectrum used for each so as to obtain half of the projections with a low energy spectrum and half with a high energy one. This revolutionary method to perform CEBT would result in a lower radiation dose and in eliminated mis-registration artifacts compared to the currently investigated methods.

6.5 Conclusion

The spectral reconstruction algorithm described here exhibits the potential to improve the image quality of DBT acquisitions over mono-energetic reconstruction methods, and to allow for dual spectrum DBT imaging with minimal impact to image quality but with a drastic reduction in dose. If these gains are confirmed in large-scale observer studies, acquisition of DBT images with alternating spectra would be greatly beneficial for a screening technology.

CHAPTER 7

X-RAY SCATTER MAP 2DPCA MODEL

In this chapter I present a study to develop a DBT X-ray scatter map model based on two dimensional principal component analysis (2DPCA) for use in the scatter correction algorithm described in chapter 2. This algorithm was expanded in chapter 5 by the replacement of patient specific Monte Carlo (MC) simulations with a library of scatter maps. The development of a scatter map model eliminates the high capacity storage requirements of a library and may reduce the initial requirement of large computing time to generate the scatter map library.

The ability of the 2DPCA model to generate novel scatter maps was tested by the generation of a scatter map that is not part of the analysis. Comparison of the modeled scatter map with the MC simulated scatter map demonstrates that the model exhibits the potential to serve as the X-ray scatter map generation method for the scatter correction algorithm.

7.1 Significance

As previously stated in this thesis, current tomosynthesis systems, including the Selenia Dimensions (Hologic Inc., Bedford, MA) used in this study, lack X-ray scatter reduction measures, be it in software or hardware. This leads to the inclusion of the entirety of the X-ray scatter signal in the tomosynthesis projections, which is expressed in the higher ratio of the number of scattered X-rays incident on the detector to that of the number of incident primary (non-scattered) X-rays, or the scatter-to-primary ratio (SPR). I have previously described a software-based X-ray scatter correction algorithm that is applied to the acquired tomosynthesis projections before reconstruction in chapter 2 that

has shown great potential in improving image quality. In chapter 5, this algorithm was expanded by incorporating an X-ray scatter map library.

Here I present a study on the incorporation of a model of X-ray scatter into the scatter correction algorithm, both to reduce computation time and to eliminate the need for the high capacity storage of a scatter map library. This model is based on a two dimensional principal component analysis (2DPCA) of a database of X-ray scatter maps of simulated breast volumes of differing shape and size.

The 2DPCA model demonstrates the potential to act as a substitute for the library of X-ray scatter maps in the scatter correction algorithm. With this improvement, the optimized scatter correction algorithm is complete and ready large scale clinical evaluation.

7.2 Materials and Methods

The 2DPCA-based model was developed in three steps. First, a library of cranio-caudal (CC) view DBT scatter maps was assembled. Then 2DPCA was performed on these scatter maps to condense the information in the images down to a reduced number of linearly independent variable vectors. And finally, testing of the 2DPCA model's ability to estimate scatter maps that were not part of the library was performed.

7.2.1 X-ray Scatter Map Library

X-ray scatter maps of 3D breast volumes (3DV) undergoing DBT were computed to form a scatter map library. The 3DVs are built using the 2D breast shape (2DS) generated by the 4-component PCA model from chapter 4 and an estimated curvature of the compressed breast volume. A script was written in IDL 8.0 (Exelis VIS, Boulder, CO) that generates a binary image stack that represents the modeled breast volume with a given compressed thickness. The process of creating a 3DV from PCA model-generated 2DS is fully described in section 5.2.1.

Each 3DV can thus be described by the four principal components α, β, γ, δ, and the compressed breast thickness *Th*. The scatter map library used in this study consists of 3DVs characterized by the principal component values in Table 7.1. The mean (μ) and standard deviation (σ) of each principal component, as detailed in chapter 4, was used for these selections. As the first component, α, has been shown to be the greatest determinant of the breast shape and size, 5 values of α were chosen. For the other three components, only 3 values were chosen. *Th* was chosen to span the range from 2 cm to 8 cm, in 2 cm steps. Thus, a total of 540 distinct 3DVs were generated from the PCA model of breast volume undergoing DBT acquisition.

Table 7.3 – Selected Values for Principal Components of the Breast Shape – To determine the input values to the PCA breast shape model, the mean (μ) and standard deviation (σ) of each component were used.

Table 7.1 Selected Values For Principal Components of Library

Principal Component	Number of Values	μ	σ	Selected Values
α	5	-3.1852	15.7792	μ + 2σ μ + σ μ μ - σ μ - 2σ
β	3	-0.2959	8.2254	μ + 2σ μ μ - 2σ
γ	3	-0.0654	3.6579	μ + 2σ μ μ - 2σ
δ	3	-0.1286	1.9939	μ + 2σ μ μ - 2σ

Subsequently, a library of 540 X-ray scatter map sets was generated using the X-ray scatter MC simulation described in chapter 5, which is a voxelized version of that used by Feng et al.[109] The simulation is implemented in C++ and is based on the Geant4 Monte Carlo simulation toolkit,[54, 55] and matches the acquisition geometry of the Selenia

Dimensions system for the CC view. The image detector measures 24 x 29 cm, with a pixel pitch of 70 μm, but since it is known that X-ray scatter varies slowly over the detector, a pixel pitch of 7.0 mm was used in the MC simulations to decrease the simulation time by reducing the noise in the results. The glandular fraction of the breast for all simulations was set to 14.3%, which has previously been shown to be the population average.[75] This is a reasonable assumption as it has been shown that scatter varies to a minimal extent with glandular fraction.[20]

7.2.2 Two Dimensional Principal Component Analysis

Each X-ray scatter map from the library can be defined as a *m* x *n* matrix of correlated observations which we can reduce using 2DPCA to a smaller set of linearly independent vectors of variables that encapsulates most of the information contained within the datapoints. 2D PCA has previously been used in imaging applications such as facial recognition.[140, 141] To perform the analysis, the *m* x (*n+6)* matrix $\boldsymbol{X_i}$ is formed by concatenating to each scatter map *i* a *m* x 6 matrix $\boldsymbol{P_i}$ of the parameter vector $\boldsymbol{p_i}$, which is comprised of the four principal components α, β, γ, δ, the compressed breast thickness *Th*, and acquisition angle *θ*. This is done to associate the 3DV parameters and *θ* to each scatter map.

$$\boldsymbol{P_i} = \begin{bmatrix} \alpha & \cdots & \alpha \\ \beta & \cdots & \beta \\ \gamma & \cdots & \gamma \\ \delta & \cdots & \delta \\ Th & \cdots & Th \\ \theta & \cdots & \theta \end{bmatrix} \qquad \text{(Equation 7.1)}$$

Then (*n+6*) x (*n+6*) covariance matrix, cov(***X***), is then formed from the mean-centered ***X*** matrices:

$$cov(\boldsymbol{X}) = \frac{\sum_{i=1}^{N}(\boldsymbol{X_i} - \bar{\boldsymbol{X}})^T(\boldsymbol{X_i} - \bar{\boldsymbol{X}})}{N-1} \qquad \text{(Equation 7.2)}$$

where N is the total number of images processed and $\bar{\boldsymbol{X}}$ is the mean of the $\boldsymbol{X}$ matrices. The principal components of the dataset are then represented by the eigenvectors $\boldsymbol{e_j}$, which can be obtained by solving the eigenvalue problem:

$$cov(\boldsymbol{X})\boldsymbol{e_j} = v_j\boldsymbol{e_j} \quad \text{(Equation 7.3)}$$

where the associated eigenvalues (v_j) also represent the variance of each principal component. The k principal components with the greatest v_j were selected and their respective $\boldsymbol{e_j}$ assembled into a n x k principal component matrix $\boldsymbol{U}$. This matrix can be used to describe any of the scatter maps that were used in the 2DPCA, and forms the basis of the scatter map model. The m x k image feature matrix $\boldsymbol{V_i}$ for image i can be calculated as follows:

$$\boldsymbol{V_i} = \boldsymbol{X_i}\boldsymbol{U} \quad \text{(Equation 7.4)}$$

Initial validation of the model was done by comparing the mean absolute error (MAE) between each scatter map from the library and its corresponding estimate from the 2DPCA model. The estimate is obtained by first calculating $\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}$ as follows:

$$\widetilde{\boldsymbol{X}}_{\boldsymbol{i}} = \boldsymbol{V_i}\boldsymbol{U^T} \quad \text{(Equation 7.5)}$$

Following this calculation, the scatter map estimate and estimated parameter vector $\widetilde{\boldsymbol{p}}_{\boldsymbol{i}}$ are extracted from $\widetilde{\boldsymbol{X}}_{\boldsymbol{i}}$.

All 2DPCA analysis was performed using MATLAB® R2011b (MathWorks, Natick, MA).

7.2.3 X-Ray Scatter Map Model

Generation of a novel scatter map from the model is initiated by the selection of the target parameter vector $\boldsymbol{t}$, which consists of the target 3DV parameters [α, β, γ, δ, Th] and θ. The closest matching scatter map from the library is identified by searching for the minimum root mean square error (RMSE) between $\boldsymbol{t}$ and the parameter vectors in the library. The image feature matrix $\boldsymbol{V_j}$ of this scatter map j, calculated using Equation 7.4, is then chosen for the initial guess to a targeted brute force search.

In each iteration, the search varies the image feature matrix $\boldsymbol{V_{test}}$ by individually varying each of the *k* image feature vectors that comprise it. We can then calculate the (*m+6)* x *n* matrix $\boldsymbol{Y_i}$ as follows:

$$\boldsymbol{Y_i} = \boldsymbol{V_{test}U^T} \quad \text{(Equation 7.6)}$$

Following this calculation, the resulting parameter vector $\boldsymbol{p_j}$ is extracted from $\boldsymbol{Y_i}$, and compared to $\boldsymbol{t}$. The output scatter map is extracted from the $\boldsymbol{Y_i}$ that exhibits the minimum RMSE between $\boldsymbol{p_j}$ and $\boldsymbol{t}$.

Testing of the model was performed by the generation of two novel 3DVs that are not part of the scatter map library, and performing MC simulations to obtain the scatter map sets of the test 3DVs. The test 3DV parameters are then used to initiate a search from the 2DPCA model, and finally the true scatter maps are compared to the 2DPCA model output by comparing the signal profiles and by calculating the MAE between the maps.

7.1 Results

7.3.1 Two Dimensional Principal Component Analysis

The 2DPCA of the X-ray scatter maps revealed that 98.7% of the variance is contained in the first six principal components (Table 7.2).

Table 7.2 – Percentage of total variance contained in the first six principal components of the 2D Principal Component Analysis – Shown here is the variance contained within each individual component and the cumulative variance contained within each component and its predecessors.

Table 7.2 Percentage of Total Variance of 2DPCA

Principal Component #	Individual Component/Cumulative (%)
1	50.1/50.1
2	22.7/72.8
3	15.2/88.0
4	5.6/93.6
5	3.4/97.0
6	1.7/98.7

However, during the initial validation testing it was found that a minimum of 15 principal components were required to recapture the scatter maps from the library without artifacts. Using a selection of k=15 principal component vectors for the $\boldsymbol{U}$ matrix, validation testing showed that the average MAE between each scatter map from the library and its corresponding estimate from the 2DPCA model is 0.0024% (SE =0.00033%).

7.3.2 X-Ray Scatter Map Model

The 2DPCA model successfully generated novel X-ray scatter maps of two 3DVs that were not part of the library (Table 7.3), and therefore not used in the analysis . This process required less than 30 minutes for each 3DV. Visual examination reveals that the modeled scatter maps are very similar in appearance to the MC simulated scatter maps (Figure 7.1A-B, Figure 7.2A-B). In addition, signal level profiles drawn through the center of the scatter maps appears to be very similar up to the point furthest from the

chest wall where X-rays pass through breast tissue (i.e. not open field), as can be seen in Figure 7.1C and in Figure 7.2C. Since these maps will be used to reduce the scatter signal in the breast projection before reconstruction, any error in the scatter map generation in the open field is not of concern. The MAEs inside the breast shapes of the novel X-ray scatter maps are 2.2% and 2.5%.

Table 7.3 –3D Breast Test Volumes – Comparison of the image quality of the spherical embedded masses in the phantom between the spectral reconstruction (SR) images and the maximum likelihood estimation maximization (MLEM) images.

Table 7.3 3D Breast Test Volumes

Parameter	Test Volume A	Test Volume B
α	$\mu + 1.5\sigma$	$\mu + \sigma$
β	$\mu + 2\sigma$	$\mu + \sigma$
γ	$\mu - 2\sigma$	$\mu - \sigma$
δ	μ	$\mu + 2\sigma$
Th	4 cm	6 cm
θ	-0.003 deg	-0.003 deg

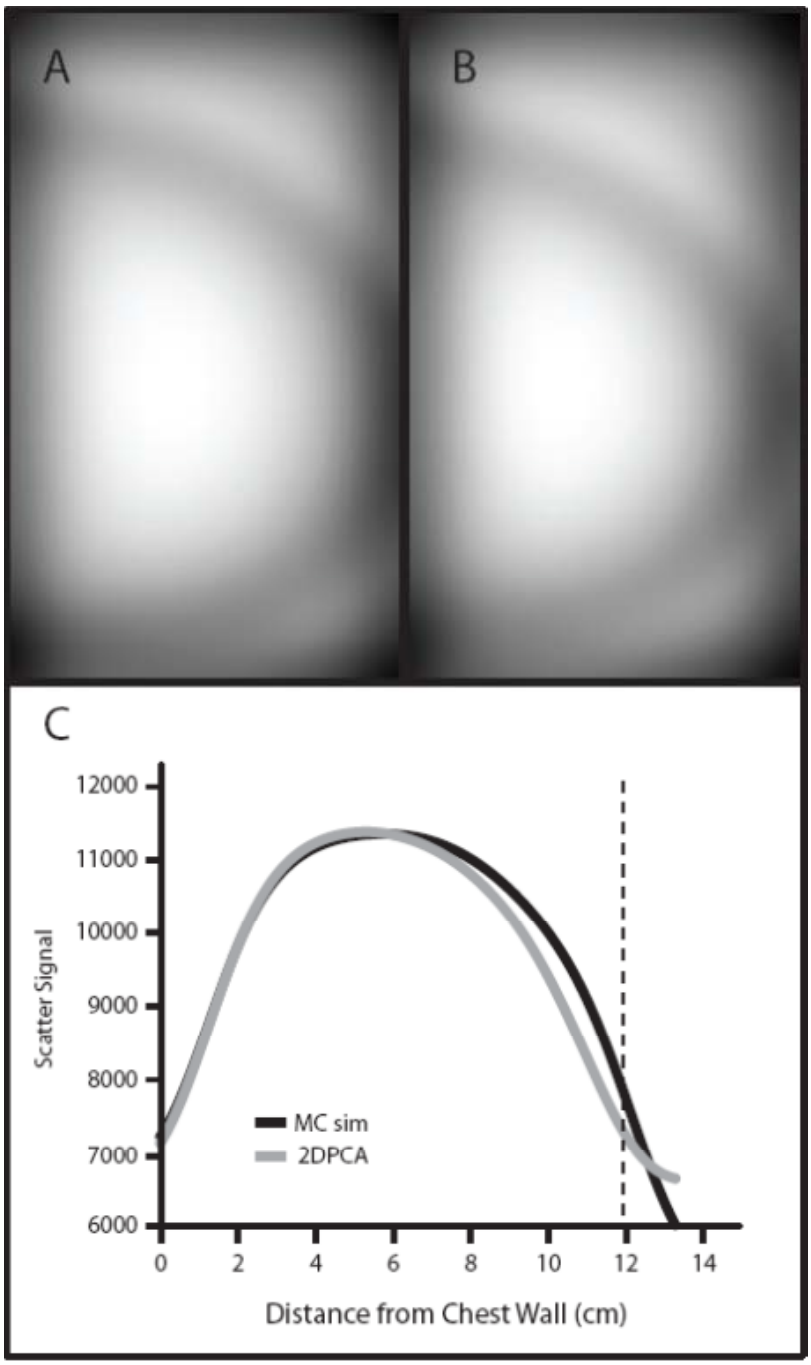

Figure 7.1 – Novel Scatter Map of Test Volume A – An X-ray scatter map of Test Volume A (delineated in Table 7.3) is shown in (A) and the corresponding scatter map generated by the 2DPCA model is shown in (B). The MAE inside the breast shape of the modeled scatter map is 2.2%. The graph in (C) depicts the signal profile through the center of the both the Monte Carlo simulated (black) scatter map and the corresponding scatter map generated by the 2DPCA model (grey). The profiles are very similar inside the breast tissue up to the furthest distance from the chest wall (dashed black line).

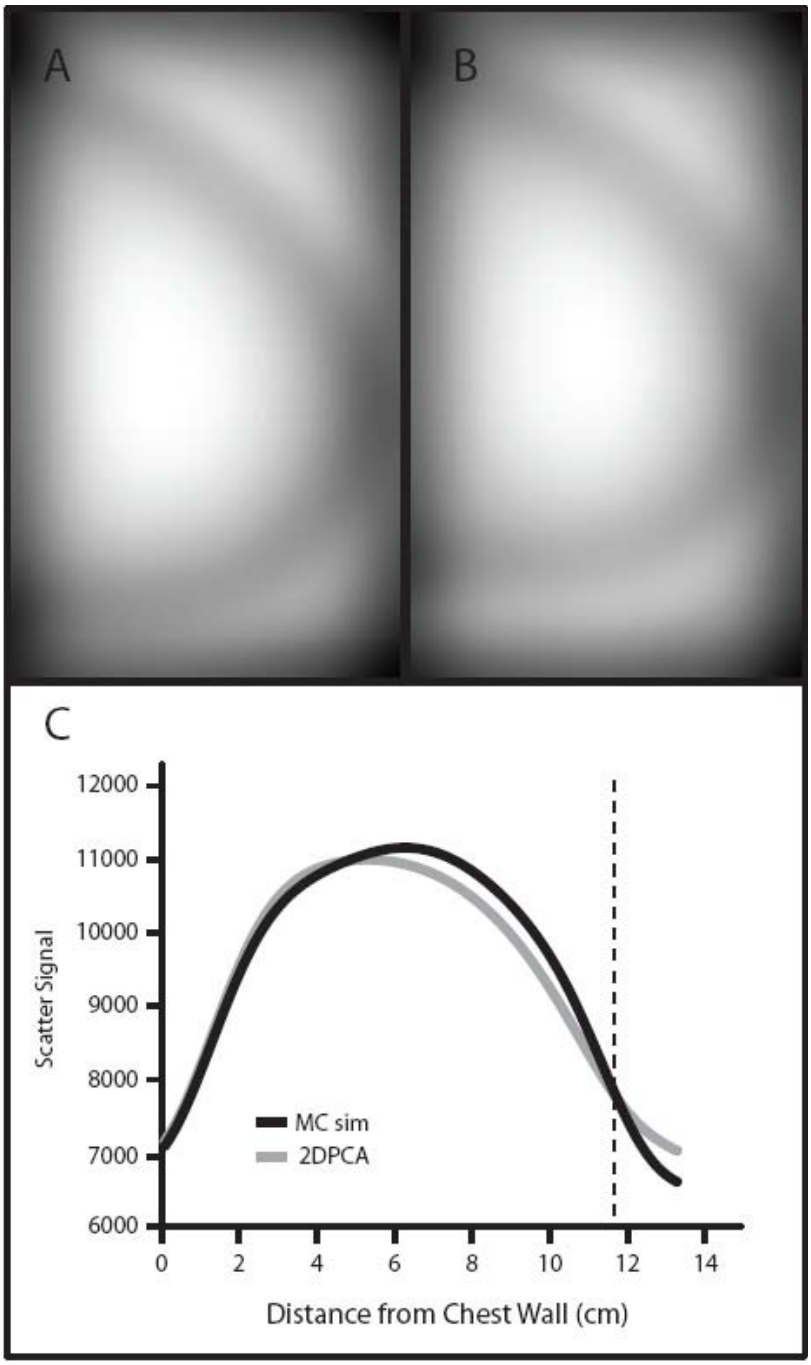

Figure 7.3 – Novel Scatter Map of Test Volume B – An X-ray scatter map of Test Volume B (delineated in Table 7.3) is shown in (A) and the corresponding scatter map generated by the 2DPCA model is shown in (B). The MAE inside the breast shape of the modeled scatter map is 2.5%. The graph in (C) depicts the signal profile through the center of the both the Monte Carlo simulated (black) scatter map and the corresponding scatter map generated by the 2DPCA model (grey). The profiles are very similar inside the breast tissue up to the furthest distance from the chest wall (dashed black line).

7.4 Discussion

The study presented in this chapter shows that a 2DPCA model of X-ray scatter maps exhibits the potential to predict the x-ray scatter signal distribution in DBT imaging. Following the incorporation of this model into the DBT scatter correction algorithm, in a hypothetical clinical setting, the central angle projection of the patient image is subjected to the automated edge detection (section 4.2.2) and the PCA (section 4.2.3) previously described to obtain the patient 3DV parameters α, β, γ, δ. The readout from the DBT system provides *Th*, and the 2DPCA model described here can then be used to generate the X-ray scatter map set for use in the scatter correction algorithm described in chapter 2. This CC view model requires less than 1 GB of data storage, as compared 120 gigabytes for the scatter map library presented in chapter 5. There is also a one-time initial computation requirement of 210,000 hours associated with the library. In addition, a medio-lateral oblique (MLO) view 2DPCA model would have similar requirements to the model presented here, while a MLO view scatter map library could more than double the computation time and storage requirements of the library presented in chapter 5.

Currently, the 2DPCA model is built upon a library of 540 scatter map sets, which is the same library created in section 5.2.2. This library was found to be insufficient in terms of range and granularity to cover the majority of breast shapes and sizes encountered clinically, as detailed in section 5.3.1. The required expansion of that preliminary scatter map library,, suggests that the 2DPCA model described here could also benefit from an expanded library. Work is currently underway to do so by including mammograms and central angle DBT projections acquired by the Selenia Dimensions system.

Finally, this study focused on CC view DBT images. A similar study would need to be performed to accommodate the MLO view. A PCA model of breast shapes undergoing MLO mammographic imaging is already complete, and is detailed in chapter

4. A MLO view X-ray scatter map library, similar to the CC view library detailed in chapter 5, would need to be built, tested and then analyzed following the method described in section 5.2.3 to generate a second 2DPCA model.

The 2PDCA model could also benefit from the incorporation of an improved search algorithm in terms of both efficiency and accuracy. The current targeted brute force method searches only the parameter space close to the closest matching scatter map in the library. A gradient descent search algorithm was tested but did not converge successfully. Ongoing work to incorporate a brute force approach with heuristics could potentially eliminate large portions of the parameter space to search.

7.5 Conclusion

The 2DPCA scatter map model presented in this chapter exhibits the potential to improve the SC algorithm by eliminating the computing time needed for patient specific MC simulations and the high storage space requirements of a X-ray scatter map library, as detailed in chapter 5.

CHAPTER 8

CONCLUSION

This thesis documents a multi-faceted, software-based approach to enhance the image quality of DBT imaging. The successful completion of the development, optimization, and testing of the scatter correction (SC) and spectral reconstruction (SR) algorithms result in improved DBT image quality. They are both ready for implementation and clinical evaluation in a DBT system, and have the potential to increase the sensitivity and specificity of breast cancer detection and diagnosis with breast tomosynthesis. In addition, a comprehensive dosimetric characterization of the Selenia Dimensions (Hologic Inc., Bedford, MA) DBT system has also been performed, and the feasibility of a new dual-spectrum, single-acquisition, DBT imaging technique has also been evaluated.

The software-based SC algorithm presented in chapter 2 was successfully implemented and tested on DBT acquisitions of both phantoms and patients. Measurable improvements in the image quality of the phantoms images were found and the application of this method to patient images also demonstrated improvement in image quality. The incorporation of the cranio-caudal (CC) view principal component analysis (PCA) model of breast shape, presented in chapter 4, into the SC algorithm has improved the efficiency of the algorithm and provided the size and range required of a pre-computed CC view X-ray scatter map library, as described in chapter 5. The observer study, also detailed in chapter 5 is also the first test of the SC algorithm with patient images and human observers, and demonstrates the potential it has to improve the clinical performance of DBT. Subsequently, the 2DPCA X-ray scatter map model detailed in chapter 7 exhibits the potential to further improve the SC algorithm by eliminating the storage space needed for the scatter map library.

The SR algorithm detailed in chapter 6 exhibits the potential to improve the image quality of DBT acquisitions over mono-energetic reconstruction methods, and to allow for dual spectrum DBT imaging with minimal impact to image quality but with a drastic reduction in dose. Acquisition of DBT images with alternating spectra could be greatly beneficial for a screening technology.

Beyond the SC and SR algorithms, other beneficial outcomes of the work presented in this thesis include the objective PCA models of breast shape presented in chapter 4 and the dosimetric characterization of a clinical DBT system presented in chapter 3. The PCA models of breast shape demonstrate the potential to be useful tools for various breast imaging applications, both for analyzing acquired images and for generating realistic breast shapes. The ability to characterize a breast shape using just a few parameters could be great utility in many applications. The dosimetry study is an important contribution to the field of knowledge, and has major implications for patient care. In addition, comprehension of the dosimetric characteristics of this particular imaging system in both FFDM and DBT mode are critical objectives of this dissertation, as the improvement of DBT image quality must be achieved without compromising radiation dose.

REFERENCES

1) National Breast Cancer Foundation, "Signs and symptoms," (2006).
2) American Cancer Society, "Cancer Facts & Figures 2010," (American Cancer Society, Atlanta, GA, 2010)
3) American Cancer Society, *Detailed Guide: Breast Cancer*. (2007).
4) A. Jemal, T. Murray, E. Ward, et al., "Cancer statistics, 2005," CA Cancer J Clin **55**, 10-30 (2005).
5) U. C. NCI-funded Breast Cancer Surveillance Consortium co-operative agreement (U01CA63740, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040). "Sensitivity and Specificity for 4,032,556 Screening Mammography Examinations from 1996 - 2005," November 29, 2007 [accessed May 20, 2009]; Available from: http://breastscreening.cancer.gov/data/benchmarks/screening/tableSensSpec.html
6) U. C. NCI-funded Breast Cancer Surveillance Consortium co-operative agreement (U01CA63740, U01CA86082, U01CA63736, U01CA70013, U01CA69976, U01CA63731, U01CA70040). "Abnormal Interpretations for 4,032,556 Screening Mammography Examinations from 1996 - 2005," November 29, 2007 [accessed May 20, 2009]; Available from: http://breastscreening.cancer.gov/data/benchmarks/screening/table3.html
7) R. D. Rosenberg, W. C. Hunt, M. R. Williamson, et al., "Effects of age, breast density, ethnicity, and estrogen replacement therapy on screening mammographic sensitivity and cancer stage at diagnosis: review of 183,134 screening mammograms in Albuquerque, New Mexico," Radiology **209**, 511-518 (1998).
8) E. D. Pisano, C. Gatsonis, E. Hendrick, et al., "Diagnostic performance of digital versus film mammography for breast-cancer screening," N Engl J Med **353**, 1773-1783 (2005).
9) L. Tabar, G. Fagerberg, H. H. Chen, et al., "Efficacy of breast cancer screening by age. New results from the Swedish Two-County Trial," Cancer **75**, 2507-2517 (1995).
10) A. Tingberg, I. Andersson, D. Ikeda, et al., "BIRADS Classification in Breast Tomosynthesis Compared to Mammography and Ultrasonography," in 9th International Workshop on Digital Mammography, (Tucson, AZ, 2008), p. 67-73.
11) L. T. Niklason, B. T. Christian, L. E. Niklason, et al., "Digital tomosynthesis in breast imaging," Radiology **205**, 399-406 (1997).
12) S. Suryanarayanan, A. Karellas, S. Vedantham, et al., "Comparison of tomosynthesis methods used with digital mammography," Acad. Radiol. **7**, 1085-1097 (2000).
13) S. Suryanarayanan, A. Karellas, S. Vedantham, et al., "Evaluation of linear and nonlinear tomosynthetic reconstruction methods in digital mammography," Acad. Radiol. **8**, 219-224 (2001).
14) T. Wu, Ph.D. Thesis, Brandeis University, 2002.
15) T. Wu, R. H. Moore, E. A. Rafferty, et al., "A comparison of reconstruction algorithms for breast tomosynthesis," Med. Phys. **31**, 2636-2647 (2004).

16) J. T. Rakowski and M. J. Dennis, "A comparison of reconstruction algorithms for C-arm mammography tomosynthesis," Med. Phys. **33**, 3018-3032 (2006).
17) Y. Zhang, H.-P. Chan, B. Sahiner, et al., "A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis," Med. Phys. **33**, 3781-3795 (2006).
18) J. Zhou, B. Zhao and W. Zhao, "A computer simulation platform for the optimization of a breast tomosynthesis system," Med. Phys. **34**, 1098-1109 (2007).
19) I. Sechopoulos, S. Suryanarayanan, S. Vedantham, et al., "Computation of the glandular radiation dose in digital tomosynthesis of the breast," Med. Phys. **34**, 221-232 (2007).
20) I. Sechopoulos, S. Suryanarayanan, S. Vedantham, et al., "Scatter radiation in digital tomosynthesis of the breast," Med. Phys. **34**, 564-576 (2007).
21) I. Sechopoulos and C. J. D'Orsi, "Glandular radiation dose in tomosynthesis of the breast using tungsten targets," J. Appl. Clin. Med. Phys. **9**, 161-171 (2008).
22) I. Sechopoulos and C. Ghetti, "Optimization of the acquisition geometry in digital tomosynthesis of the breast," Med. Phys. **36**, 1199-1207 (2009).
23) B. Ren, C. Ruth, J. Stein, et al., "Design and performance of the prototype full field breast tomosynthesis system with selenium based flat panel detector," Proc. SPIE **5745**, 550-561 (2005).
24) J. W. Eberhard, P. Staudinger, J. Smolenski, et al., "High-speed large-angle mammography tomosynthesis system," Proc. SPIE **6142**, 61420C-61411 (2006).
25) R. H. Moore, D. B. Kopans, E. A. Rafferty, et al., "Initial Callback Rates for Conventional and Digital Breast Tomosynthesis Mammography Comparison in the Screening Setting," in Radiological Society of North America 93rd Scientific Assembly and Annual Meeting Program, (Radiological Society of North America, Chicago, IL, 2007), p. 381.
26) M. A. Helvie, M. A. Roubidoux, L. M. Hadjiiski, et al., "Tomosynthesis Mammography versus Conventional Mammography: Comparison of Breast Masses Detection and Characterization," in Radiological Society of North America 93rd Scientific Assembly and Annual Meeting Program, (Radiological Society of North America, Chicago, IL, 2007), p. 381.
27) S. P. Poplack, T. D. Tosteson, C. A. Kogel, et al., "Digital Breast Tomosynthesis: Initial Experience in 98 Women with Abnormal Digital Screening Mammography," Am. J. Roentgenol. **189**, 616-623 (2007).
28) I. Andersson, D. Ikeda, S. Zackrisson, et al., "Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings," Eur. Radiol. **18**, 2817-2825 (2008).
29) A. Smith, E. Rafferty and L. Niklason, "Clinical Performance of Breast Tomosynthesis as a Function of Radiologist Experience Level," in 9th International Workshop on Digital Mammography, (Tucson, AZ, 2008), p. 61-66.
30) W. F. Good, G. S. Abrams, V. J. Catullo, et al., "Digital Breast Tomosynthesis: A Pilot Observer Study," Am. J. Roentgenol. **190**, 865-869 (2008).
31) M. A. Helvie, "Digital Mammography Imaging: Breast Tomosynthesis and Advanced Applications," Radiol. Clin. North Am. **48**, 917-929 (2010).

32) J. T. Dobbins, 3rd, "Tomosynthesis imaging: At a translational crossroads," Med. Phys. **36**, 1956-1967 (2009).
33) I. Sechopoulos, "A review of breast tomosynthesis. Part I. The image acquisition process," Medical Physics **40** (2013).
34) X. Qian, R. Rajaram, X. Calderon-Colon, et al., "Design and characterization of a spatially distributed multibeam field emission x-ray source for stationary digital breast tomosynthesis," Medical Physics **36**, 4389-4399 (2009).
35) F. Sprenger, X. Calderon, E. Gidcumb, et al., "Stationary digital breast tomosynthesis with distributed field emission X-ray tube," Proceedings of SPIE **7961**, 79615I (2011).
36) F. Sprenger, X. Calderon-Colon, Y. Cheng, et al., "Distributed source x-ray tube technology for tomosynthesis imaging," Proceedings of SPIE **7622**, 76225M (2010).
37) G. Yang, X. Qian, T. Phan, et al., "Design and feasibility studies of a stationary digital breast tomosynthesis system," Nuclear Instruments and Methods in Physics Research. Section A. **648**, S220-S223 (2011).
38) X. Qian, A. Tucker, E. Gidcumb, et al., "A stationary digital breast tomosynthesis scanner," Proceedings of SPIE **8313**, 831352-831311 (2012).
39) X. Qian, A. Tucker, E. Gidcumb, et al., "High resolution stationary digital breast tomosynthesis using distributed carbon nanotube x-ray source array," Medical Physics **39**, 2090-2099 (2012).
40) A. Tucker, X. Qian, E. Gidcumb, et al., "Optimizing configuration parameters of a stationary digital breast tomosynthesis system based on carbon nanotube x-ray sources," Proceedings of SPIE **8313**, 831307-831310 (2012).
41) M. Aslund, B. Cederstrom, M. Lundqvist, et al., "Scatter rejection in multislit digital mammography," Medical Physics **33**, 933-940 (2006).
42) F. F. Schmitzberger, E. M. Fallenberg, R. Lawaczeck, et al., "Development of low-dose photon-counting contrast-enhanced tomosynthesis with spectral imaging," Radiology **259**, 558-564 (2011).
43) M. G. Wallis, E. Moa, F. Zanca, et al., "Two-View and Single-View Tomosynthesis versus Full-Field Digital Mammography: High-Resolution X-Ray Imaging Observer Study," Radiology **262**, 788-796 (2012).
44) G. M. Stevens, R. L. Birdwell, C. F. Beaulieu, et al., "Circular tomosynthesis: potential in imaging of breast and upper cervical spine--preliminary phantom and in vitro study," Radiology **228**, 569-575 (2003).
45) K. Zeng, H. Yu, S. Zhao, et al., "Digital tomosynthesis aided by low-resolution exact computed tomography," Journal of Computer Assisted Tomography **31**, 976-983 (2007).
46) D. Xia, S. Cho, J. Bian, et al., "Tomosynthesis with source positions distributed over a surface," Proc. SPIE **6913**, 69132A-69137 (2008).
47) J. Zhang and C. Yu, "A novel solid-angle tomosynthesis (SAT) scanning scheme," Med. Phys. **37**, 4186-4192 (2010).
48) G. Wu, J. G. Mainprize, J. M. Boone, et al., "Evaluation of scatter effects on image quality for breast tomosynthesis," Proc. SPIE **6510**, 65101T-65109 (2007).
49) G. Wu, J. G. Mainprize, J. M. Boone, et al., "Evaluation of scatter effects on image quality for breast tomosynthesis," Med. Phys. **36**, 4425-4432 (2009).

50) B. Liu, T. Wu, R. H. Moore, et al., "Monte Carlo simulation of x-ray scatter based on patient model from digital breast tomosynthesis," Proc. SPIE **6142**, 61421N-61429 (2006).
51) J. Chung, J. G. Nagy and I. Sechopoulos, "Numerical Algorithms for Polyenergetic Digital Breast Tomosynthesis Reconstruction," SIAM Journal on Imaging Sciences **3**, 133-152 (2010).
52) V. M. Bustamante, J. G. Nagy, S. S. Feng, et al., "Iterative Breast Tomosynthesis Image Reconstruction," SIAM Journal on Scientific Computing, in press (2013).
53) T. Wu, J. Zhang, R. Moore, et al., "Digital tomosynthesis mammography using a parallel maximum-likelihood reconstruction method," Proc. SPIE **5368**, 1-11 (2004).
54) S. Agostinelli, J. Allison, K. Amako, et al., "Geant4 - A simulation toolkit," Nucl. Instrum. Meth. A **506**, 250-303 (2003).
55) J. Allison, K. Amako, J. Apostolakis, et al., "Geant4 Developments and Applications," IEEE Trans. Nucl. Sci. **53**, 270-278 (2006).
56) B. R. Ren, C. Ruth, T. Wu, et al., "A new generation FFDM/tomosynthesis fusion system with selenium detector," Medical Imaging 2010: Physics of Medical Imaging **7622**, -1816 (2010).
57) J. M. Boone, T. R. Fewell and R. J. Jennings, "Molybdenum, rhodium, and tungsten anode spectral models using interpolating polynomials with application to mammography," Medical Physics **24**, 1863-1874 (1997).
58) A. P. Colijn and F. J. Beekman, "Accelerated simulation of cone beam X-ray scatter projections," IEEE Trans. Med. Imaging **23**, 584-590 (2004).
59) R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," Med. Phys. **12**, 252-255 (1985).
60) J. M. Boone and A. E. Chavez, "Comparison of x-ray cross sections for diagnostic and therapeutic medical physics," Med Phys **23**, 1997-2005 (1996).
61) R. C. Gonzalez and R. E. Woods, *Digital image processing*, 2nd ed. (Prentice Hall, Upper Saddle River, N.J., 2002).
62) T. Pavlidis, *Algorithms for graphics and image processing*. (Computer Science Press, Rockville, MD, 1982).
63) R. Chandrasekhar and Y. Attikiouzel, "A simple method for automatically locating the nipple on mammograms," IEEE Trans. Med. Imaging **16**, 483-494 (1997).
64) F. L. Bookstein, "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations," IEEE Trans. Pattern Anal. Mach. Intell. **11**, 567-585 (1989).
65) R. C. Gonzalez and R. E. Woods, *Digital image processing*, 3rd ed. (Prentice Hall, Upper Saddle River, N.J., 2008).
66) I. S. Kyprianou, A. Badano, B. D. Gallas, et al., "A method to estimate the point response function of digital x-ray detectors from edge measurements," Medical Imaging 2007: Physics of Medical Imaging, Pts 1-3 **6510**, U104-U115 2275 (2007).
67) H. Fujita, D. Y. Tsai, T. Itoh, et al., "A Simple Method for Determining the Modulation Transfer-Function in Digital Radiography," IEEE Trans. Med. Imag. **11**, 34-39 (1992).

68) D. Förnvik, I. Andersson, T. Svahn, et al., "The effect of reduced breast compression in breast tomosynthesis: human observer study using clinical cases," Radiat. Prot. Dosimetry **139**, 118-123 (2010).
69) R. S. Saunders, E. Samei, J. Y. Lo, et al., "Can Compression Be Reduced for Breast Tomosynthesis? Monte Carlo Study on Mass and Microcalcification Conspicuity in Tomosynthesis," Radiology **251**, 673-682 (2009).
70) A. K. W. Ma, D. G. Darambara, A. Stewart, et al., "Mean glandular dose estimation using MCNPX for a digital breast tomosynthesis system with tungsten/aluminum and tungsten/aluminum plus silver x-ray anode-filter combinations," Medical Physics **35**, 5278-5289 (2008).
71) D. R. Dance, K. C. Young and R. E. van Engen, "Estimation of mean glandular dose for breast tomosynthesis: factors for use with the UK, European and IAEA breast dosimetry protocols," Phys. Med. Biol. **56**, 453 (2011).
72) B. Ren, C. Ruth, T. Wu, et al., "A new generation FFDM/tomosynthesis fusion system with selenium detector," Proc. SPIE **7622**, 76220B-76211 (2010).
73) M. Saito, "Dual-energy approach to contrast-enhanced mammography using the balanced filter method: spectral optimization and preliminary phantom measurement," Med. Phys. **34**, 4236-4246 (2007).
74) I. Sechopoulos, S. S. J. Feng and C. J. D'Orsi, "Dosimetric characterization of a dedicated breast computed tomography clinical prototype," Med. Phys. **37**, 4110-4120 (2010).
75) M. J. Yaffe, J. M. Boone, N. Packard, et al., "The myth of the 50-50 breast," Med. Phys. **36**, 5437-5443 (2009).
76) J. M. Boone, "Glandular breast dose for monoenergetic and high-energy X-ray beams: Monte Carlo assessment," Radiology **213**, 23-37 (1999).
77) L. Wilkinson and J. C. P. Heggie. "Glandular Breast Dose: Potential Errors [eletter]," 2000 [accessed November 21, 2005]; Available from: http://radiology.rsnajnls.org/cgi/eletters/213/1/23
78) J. M. Boone, "Normalized glandular dose (DgN) coefficients for arbitrary X-ray spectra in mammography: Computer-fit values of Monte Carlo derived data," Med. Phys. **29**, 869-875 (2002).
79) G. R. Hammerstein, D. W. Miller, D. R. White, et al., "Absorbed radiation dose in mammography," Radiology **130**, 485-491 (1979).
80) J. M. Boone, K. K. Lindfors, V. N. Cooper, 3rd, et al., "Scatter/primary in mammography: Comprehensive results," Med. Phys. **27**, 2408-2416 (2000).
81) R. E. Hendrick, E. D. Pisano, A. Averbukh, et al., "Comparison of Acquisition Parameters and Breast Dose in Digital Mammography and Screen-Film Mammography in the American College of Radiology Imaging Network Digital Mammographic Imaging Screening Trial," American Journal of Roentgenology **194**, 362-369 (2010).
82) M. J. Michell, R. K. Wasan, A. Iqbal, et al., "Two-view 2D digital mammography versus one-view digital breast tomosynthesis," in Royal College of Radiologists Breast Group Annual Scientific Meeting 2010, (BioMed Central, Brighton, UK, 2010), p. 3.
83) E. A. Rafferty, "Digital mammography: Novel applications," Radiologic Clinics of North America **45**, 831 (2007).

84) D. Gur, G. S. Abrams, D. M. Chough, et al., "Digital Breast Tomosynthesis: Observer Performance Study," American Journal of Roentgenology **193**, 586-591 (2009).
85) N. Jamal, K. H. Ng and D. McLean, "A study of mean glandular dose during diagnostic mammography in Malaysia and some of the factors affecting it," Br. J. Radiol. **76**, 238-245 (2003).
86) K. C. Young and A. Burch, "Radiation doses received in the UK Breast Screening Programme in 1997 and 1998," Br. J. Radiol. **73**, 278-287 (2000).
87) American Cancer Society, "Breast Cancer Facts and Figures 2011-2012," (American Cancer Society, Atlanta, GA, 2012)
88) NCI-funded Breast Cancer Surveillance Consortium co-operative agreement. "Performance Measures for 1,960,150 Screening Mammography Examinations from 2002 to 2006 by Age --- based on BCSC data as of 2009," November 15, 2011 [accessed April 26, 2012]; Available from: http://breastscreening.cancer.gov/data/performance/screening/2009/perf_age.html
89) P. R. Bakic, M. Albert, D. Brzakovic, et al., "Mammogram synthesis using a 3D simulation. I. Breast tissue model and image acquisition simulation," Med. Phys. **29**, 2131-2139 (2002).
90) P. R. Bakic, M. Albert, D. Brzakovic, et al., "Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture," Med. Phys. **29**, 2140-2151 (2002).
91) P. R. Bakic, M. Albert, D. Brzakovic, et al., "Mammogram synthesis using a three-dimensional simulation. III. Modeling and evaluation of the breast ductal network," Med. Phys. **30**, 1914-1925 (2003).
92) P. R. Bakic, C. Zhang and A. D. Maidment, "Development and characterization of an anthropomorphic breast software phantom based upon region-growing algorithm," Med. Phys. **38**, 3165-3176 (2011).
93) B. A. Lau, I. Reiser, R. M. Nishikawa, et al., "A statistically defined anthropomorphic software breast phantom," Med. Phys. **39**, 3375-3385 (2012).
94) I. Reiser, S. Lee and R. M. Nishikawa, "On the orientation of mammographic structure," Med. Phys. **38**, 5303-5306 (2011).
95) K. Bliznakova, I. Sechopoulos, I. Buliev, et al., "BreastSimulator: A software platform for breast x-ray imaging research," Journal of Biomedical Graphics and Computing **2**, 1-14 (2012).
96) C. M. Li, W. P. Segars, G. D. Tourassi, et al., "Methodology for generating a 3D computerized breast phantom from empirical data," Med. Phys. **36**, 3122-3131 (2009).
97) D. R. Dance, "Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose," Phys. Med. Biol. **35**, 1211-1219 (1990).
98) D. R. Dance, C. L. Skinner, K. C. Young, et al., "Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol," Phys. Med. Biol. **45**, 3225-3240 (2000).
99) X. Wu, G. T. Barnes and D. M. Tucker, "Spectral dependence of glandular tissue dose in screen-film mammography," Radiology **179**, 143-148 (1991).

100) X. Wu, E. L. Gingold, G. T. Barnes, et al., "Normalized average glandular dose in molybdenum target-rhodium filter and rhodium target-rhodium filter mammography," Radiology **193**, 83-89 (1994).
101) D. R. Dance, R. A. Hunt, P. R. Bakic, et al., "Breast dosimetry using high-resolution voxel phantoms," Radiat. Prot. Dosimetry **114**, 359-363 (2005).
102) I. Reiser and R. M. Nishikawa, "Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise," Med. Phys. **37**, 1591-1600 (2010).
103) X. Gong, S. J. Glick, B. Liu, et al., "A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam CT breast imaging," Med. Phys. **33**, 1041-1052 (2006).
104) L. Zhou, J. Oldan, P. Fisher, et al., "Low-contrast lesion detection in tomosynthetic breast imaging using a realistic breast phantom," Proc. SPIE **6142**, 61425A-61412 (2006).
105) R. A. Hunt, D. R. Dance, P. R. Bakic, et al., "Calculation of the properties of digital mammograms using a computer simulation," Radiat. Prot. Dosimetry **114**, 395-398 (2005).
106) K. Bliznakova, Z. Kolitsi and N. Pallikarakis, "Dual-energy mammography: simulation studies," Phys. Med. Biol. **51**, 4497-4515 (2006).
107) F. J. P. Richard, P. R. Bakic and A. D. A. Maidment, "Mammogram registration: a phantom-based evaluation of compressed Breast Thickness variation effects," IEEE Trans. Med. Imaging **25**, 188-197 (2006).
108) D. Kontos, C. Zhang, N. Ruiter, et al., "Evaluating the Effect of Tomosynthesis Acquisition Parameters on Image Texture: A Study Based on an Anthropomorphic Breast Tissue Software Model," in 9th International Workshop on Digital Mammography, (Springer Berlin / Heidelberg, Tucson, AZ, 2008), p. 491-498.
109) S. S. J. Feng and I. Sechopoulos, "A software-based x-ray scatter correction method for breast tomosynthesis," Medical Physics **38**, 6643-6653 (2011).
110) A. K. W. Ma, D. G. Darambara, A. Stewart, et al., "Mean glandular dose estimation using MCNPX for a digital breast tomosynthesis system with tungsten/aluminum and tungsten/aluminum + silver x-ray anode-filter combinations," Med. Phys. **35**, 5278-5289 (2008).
111) N. Buls, I. Wathion, L. Mommaerts, et al., "Development of a multi-spectral, multi-geometry computational model for X-ray breast imaging," Radiat. Prot. Dosimetry **139**, 343-347 (2010).
112) R. C. Gonzalez and R. E. Woods, *Digital image processing*, 2nd ed. (Prentice Hall, Upper Saddle River, N.J., 2002).
113) T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," Proc. SPIE **4322**, 236-248 (2001).
114) I. T. Jolliffe, *Principal component analysis*, 2nd ed. (Springer, New York, 2002).
115) D. C. Barber, "The use of principal components in the quantitative analysis of gamma camera dynamic studies," Phys. Med. Biol. **25**, 283-292 (1980).
116) F. Pedersen, M. Bergström, E. Bengtsson, et al., "Principal component analysis of dynamic positron emission tomography images," Eur. J. Nucl. Med. Mol. Imaging **21**, 1285-1292 (1994).

117) D. Rueckert, A. Frangi and J. Schnabel, "Automatic Construction of 3D Statistical Deformation Models Using Non-rigid Registration," in Medical Image Computing and Computer-Assisted Intervention – MICCAI, (Utrecht, The Netherlands, 2001), p. 77-84.
118) S. S. J. Feng and I. Sechopoulos, "Clinical Digital Breast Tomosynthesis System: Dosimetric Characterization," Radiology **263**, 35-42 (2012).
119) T. Ojala, J. Nappi and O. Nevalainen, "Accurate segmentation of the breast region from digitized mammograms," Comput Med Imag Grap **25**, 47-59 (2001).
120) C. Zhou, H. P. Chan, N. Petrick, et al., "Computerized image analysis: Estimation of breast density on mammograms," Medical Physics **28**, 1056-1069 (2001).
121) M. M. Goodsitt, H. P. Chan, B. Liu, et al., "Classification of compressed breast shapes for the design of equalization filters in x-ray mammography," Med Phys **25**, 937-948 (1998).
122) R. J. Ferrari, R. M. Rangayyan, J. E. L. Desautels, et al., "Identification of the breast boundary in mammograms using active contour models," Med. Biol. Eng. Comput. **42**, 201-208 (2004).
123) S. M. Kwok, R. Chandrasekhar, Y. Attikiouzel, et al., "Automatic pectoral muscle segmentation on mediolateral oblique view mammograms," IEEE Transactions on Medical Imaging **23**, 1129-1140 (2004).
124) E. A. Krupinski and H. Roehrig, "The influence of a perceptually linearized display on observer performance and visual search," Acad. Radiol. **7**, 8-13 (2000).
125) K. R. Peters and G. R. Ramsby, "Natural display mode for digital DICOM-conformant diagnostic imaging," Acad. Radiol. **9**, 1025-1035 (2002).
126) I. Sechopoulos, "A review of breast tomosynthesis. Part II. Image reconstruction, processing and analysis, and advanced applications," Medical Physics **40** (2013).
127) I. Kastanis, S. Arridge, A. Stewart, et al., "3D Digital Breast Tomosynthesis Using Total Variation Regularization," in 9th International Workshop on Digital Mammography, (Tucson, AZ, 2008), p. 621-627.
128) E. Y. Sidky, I. S. Reiser, R. M. Nishikawa, et al., "Preliminary study on the impact of digital breast tomosynthesis scanning angle on micro-calcification imaging," in IEEE Nuclear Science Symposium Conference Record, (Dresden, Germany, 2008), p. 4201-4204.
129) E. Y. Sidky, Y. Duchin, I. Reiser, et al., "Optimizing algorithm parameters based on a model observer detection task for image reconstruction in digital breast tomosynthesis," in IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), (Valencia, Spain, 2011), p. 4230-4232.
130) E. Y. Sidky, X. Pan, I. S. Reiser, et al., "Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms," Med. Phys. **36**, 4920-4932 (2009).
131) E. Y. Sidky, I. Reiser, R. M. Nishikawa, et al., "Practical iterative image reconstruction in digital breast tomosynthesis by non-convex TpV optimization," Proc. SPIE **6913**, 691328-691321 (2008).
132) D. Van de Sompel and M. Brady, "Regularising limited view tomography using anatomical reference images and information theoretic similarity metrics," Med. Image Anal. **16**, 278-300 (2012).

133) D. Van de Sompel and S. M. Brady, "Robust joint entropy regularization of limited view transmission tomography using Gaussian approximations to the joint histogram," in Proceedings of the International Conference on Information Processing in Medical Imaging, (Williamsburg, VA, 2009), p. 638-650.
134) Y. Lu, H.-P. Chan, J. A. Fessler, et al., "Adaptive Diffusion Regularization for Enhancement of Microcalcifications in Digital Breast Tomosynthesis (DBT) Reconstruction," Proc. SPIE **7961**, 796117-796111 (2011).
135) M. Das, H. C. Gifford, J. M. O'Connor, et al., "Penalized Maximum Likelihood Reconstruction for Improved Microcalcification Detection in Breast Tomosynthesis," IEEE Trans. Med. Imaging **30**, 904-914 (2011).
136) A. Jerebko, M. Kowarschik and T. Mertelmeier, "Regularization Parameter Selection in Maximum a Posteriori Iterative Reconstruction for Digital Breast Tomosynthesis Digital Mammography," in 10th International Workshop on Digital Mammography, (Springer Berlin / Heidelberg, Girona, Spain, 2010), p. 548-555.
137) S. Xu and Y. Chen, "A Compton scattering suppression based image reconstruction method for Digital Breast Tomosynthesis," in IEEE International Workshop on Genomic Signal Processing and Statistics (GENSIPS), (San Antonio, TX, 2011), p. 190-193.
138) M. L. Hill, J. G. Mainprize, S. Puong, et al., "Impact of image acquisition timing on image quality for dual energy contrast-enhanced breast tomosynthesis," Proceedings of SPIE **8313**, 831308-831309 (2012).
139) A. K. Carton, S. C. Gavenonis, J. A. Currivan, et al., "Dual-energy contrast-enhanced digital breast tomosynthesis--a feasibility study," The British Journal of Radiology **83**, 344-350 (2010).
140) P. Ma, D. Yang, Y. X. Ge, et al., "Face recognition using two-dimensional nonnegative principal component analysis," Journal of Electronic Imaging **21**.
141) L. S. Oliveira, A. L. Koerich, M. Mansano, et al., "2D Principal Component Analysis for Face and Facial-Expression Recognition," Computing in Science & Engineering **13**, 9-13.

VITA

STEVE SI JIA FENG

Feng was born in Guangzhou, China, and grew up in New York City. There he attended Hunter College High School and then received a B.S. in Biomedical Engineering from Johns Hopkins University, Baltimore, MD in 2005 before attending GeorgiaTech. He has studied neuroscience and worked in medical imaging but remains an engineer to this day. When he manages to find time to take a break from working on his research, he enjoys running 9-ball racks, cheering for the New York Yankees and the New York Knicks, and tasting Bordeaux wines.