Stony Brook University

Improvements in Lesion Detection in X-Ray Breast Tomosynthesis and in Emission Tomography

A Dissertation Presented

by

Yihuan Lu

to

The Graduate School

in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Electrical Engineering

Stony Brook University

May 2015

Stony Brook University

The Graduate School

Yihuan Lu

We, the dissertation committee for the above candidate for the Ph.D. degree, hereby recommend acceptance of this dissertation

Gene Gindi, Dissertation Advisor
Associate Professor, Department of Electrical & Computer Engineering

Petar Djurić, Chairperson of Defense
Professor, Department of Electrical & Computer Engineering

John Murray, Associate Professor
Department of Electrical & Computer Engineering

Wei Zhao, Professor
Department of Radiology

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

Improvements in Lesion Detection in X-Ray Breast Tomosynthesis and in Emission Tomography

by
Yihuan Lu
Doctor of Philosophy
in
Electrical Engineering
Stony Brook University
2015

Cancerous lesions can be viewed using several medical imaging modalities. I address breast cancer as seen in contrast-enhanced dual-energy digital breast tomosynthesis (CE-DE-DBT) and liver imaging of liver neuroendocrine metastases as seen by SPECT (Single Photon Emission Computed Tomography). The overall goal is to enhance the ability to detect such lesions by improvements in image acquisition and data processing methods. In CE-DE-DBT the presence of scattered photons can yield image artifacts that mask the breast lesion. I present a scatter correction method based on interpolating scatter tails at the image periphery. I validate the method using a pinhole array technique and beam blocker method. CE-DE-DBT is also easily corrupted by patient motion. Acquiring the image using interleaved acquisition can aid in minimizing patient motion artifacts, but introduces new artifacts when combined with standard CE-DE-DBT processing methods. I introduce a modification of the post-acquisition processing - a "reconstruct-then-subtract" method that minimizes these artifacts. Then choice of reconstruction algorithm in the reconstruct-then-subtract method can also influence image quality. I demonstrate that the use of a statistical reconstruction technique, ordered subsets transmission tomography, can aid in visualizing small lesions. For SPECT, the collimator is the crucial element controlling the noise/resolution tradeoff in the image. The best collimator will yield the best performance on a task of detecting and localizing the lesion. I use the tools of

statistical decision theory on a performance metric based on the area under the localization receiver operating characteristic curve to optimize the collimator design, taking into account the important physical effects of collimator septal penetration and scatter by the high energy photons of the In-111 radionuclide. The optimal parallel-collimator is characterized by geometrical parameters (bore length, bore width, septal thickness) that allow a controlled number of photons from septal scatter and penetration into the image.

To my parents and wife

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ALROC	Area under the LROC curve
AROC	Area under the ROC curve
a-Se	Amorphous Selenium
CC	Craniocaudal
CE	Contrast Enhanced
CE-DE-DBT	Contrast Enhanced Dual Energy Digital Breast Tomosynthesis
CL	Correctly Localized
CNR	Contrast-to-Noise Ratio
CT	Computed Tomography
CUDA	Compute Unified Device Architecture
DBT	Digital Breast Tomosynthesis
DM	Digital Mammography
DS	Double-Sweep
FBP	Filtered Back Projection
FFT	Fast Fourier Transform
FN	False Negative
FP	False Positive
FWHM	Full Width at Half of Maximum
GPU	Graphics Processing Unit
HE	High Energy
IDC	Invasive Ductal Carcinoma
IL	Interleave
IO	Ideal Observer
kVp	Kilo voltage Peak
LE	Low Energy
LR	Likelihood Ratio
LROC	Location Receiver Operating Characteristic

MAP	Maximum *A Posteriori*
mAs	Milliamp-Second
MC	Monte Carlo
MLO	Mediolateral Oblique
MRI	Magnetic Resonance Imaging
NET	Neuroendocrine Tumors
OS	Ordered Subsets
OSTR	Ordered Subsets Transmission Reconstruction
PET	Positron Emission Tomography
PMT	Photomultiplier Tube
PSF	Point Spread Function
ROC	Receiver Operating Characteristic
RS	Reconstruction then Subtraction
SART	Simultaneous Algebraic Reconstruction Technique
SC	Scatter Correction
SDNR	Signal Difference to Noise Ratio
SPECT	Single Photon Emission Computed Tomography
SPR	Scatter-to-Primary Ratio
SPSF	Scatter Point Spread Function
SR	Subtraction then Reconstruction
TP	True Positive

Acknowledgments

I would like to express my deep thanks to my dissertation advisor, Gene Gindi for his intensive guidance and encouragement in all my dissertation work. I am grateful I chose Gene as my advisor, who taught me not only how to know things, but also to know what I don't know.

I appreciate the support from my committee members: Petar Djurić, John Murray and Wei Zhao. I thank Petar for his lectures on the statistical theories of signal processing and I thank John for his lectures on the linear system. In particular, I thank Wei for her support and advice on all the DBT part of my dissertation. Without her help and effort, my work on DBT projects would have been formidable.

Thanks also go to all the members of Medical Imaging Processing Laboratory and Dr. Zhao's laboratory: Lin, Santosh, Chintu, Yue-Houng, David, Amir, Boyu, Sarah, James, Adrian, Jann, Beverly and Rick. Specifically, I'd like to thank Boyu, David and Yue-Houng for the extremely valuable collaborative work and discussion. I also would like to thank Amir and Rick for the enlightening discussions on detector physics.

I would like to thank Dr. Zhengrong Liang and his lab members: Bowen, Yan, Hao Zhang, Hao Han, Yifan and Ming for making my daily life in Stony Brook interesting and enjoyable.

Special thanks also go to my "F1" basketball team members: Duo, Yuanhao, Huafeng, John, Guanchen, Jiahao, Yang, Jiantao, Pengfei, Feng, Hua, Hosan, Eddie and Xiang. They are the guys who made my Stony Brook basketball career extremely joyful and unforgettable.

Here, I greatly acknowledge my wife Hyang-Gi Song for her love, help and support in my life and study. Without her understanding and support, this dissertation could not have been done successfully.

Finally, I would like to thank my parents, Xiaoning Lu and Hong Peng for their support and encouragement in all my educational endeavors. I also deeply thank my uncle and aunt, Bruce Long Peng and Jean Ann, for their consistent support and help with my life and study.

Chapter 1 Introduction

In this chapter, I give a very simple qualitative overview of two medical imaging systems I will deal with, Digital Breast Tomosynthesis and Single Photon Emission Computed Tomography. I also give a very simple introduction to the nature of the research I have done with these systems.

1.1 Introduction

One big goal of engineers' work in medical image is to improve the images by changing the hardware or the processing software. Many medical image systems in modern radiology are tomographic. Data is collected by hardware from several angles in the form of line integrals and the underlying object reconstructed by software from this data. CT (computed tomography) is a familiar example. So an example of a form of software change that could yield better images is the use of a different reconstruction algorithm. Improved detectors could yield better energy resolution yielding better images in some forms of nuclear medical imaging.

One needs to quantify what is meant by a "better" image. We usually have two types of definitions of medical image quality. One is usually simple, which can be directly acquired from one single image. Examples are contrast and forms of contrast-to-noise ratio (CNR). We define another definition of image quality according to how well it supports performance on a relevant medical task. For example, if a new system with one combination of hardware/ software processing yields an improved ability to detect subtle lesions relative to an older system, then the newer system is "improved". The observer detecting the lesion is usually a human - the radiologist - but for engineering studies, it is often practical to replace the human observer by a mathematical observer. We shall talk in more detail about medical tasks and associated observers.

In my research work I am dealing with two tomographic "modalities" or forms of imaging. One is digital breast tomosynthesis (DBT). In DBT, one tries to obtain a 3-D tomographic image of the breast by reconstructing detected X-ray irradiation from a range of angles. The other is single photon emission computed tomography (SPECT). Here a radiopharmaceutical is injected

into the patient and a reconstruction is used to reconstruct the radionuclide density within the body from the γ-rays collected from many angles by a gamma camera. The familiar modality of PET is similar to SPECT but has some fundamental differences in γ -ray emission.

For DBT, my work involves making specific improvements in data processing for a particular form of DBT called CE-DE-DBT, contrast-enhanced dual-energy DBT. In particular, these improvements involve (i) the reconstruction step and (ii) the removal of scattered X-rays that degrade the image quality. For SPECT, my work centers on task performance evaluation of collimators using a novel task in which the usual detection task is replaced by one that is more clinically realistic and requires the observer to localize (search for) and detect a lesion. The collimator optimization involves complicated physics modeling of effects that can affect performance.

1.2 Basics of X-ray breast imaging and breast tomosynthesis

Here we describe some basic concepts in X-ray mammography. X-ray mammography acquires breast images from one angle rather than an angular range as in DBT, but it is useful to describe it before describing DBT. We focus on basics and leave description of details of the particular modality - Contrast Enhanced Dual Energy Digital Breast Tomosynthesis (CE-DE-DBT) for Chapter 2.

Breast cancer is the most prevalent cancer in women. Among 10 most commonly diagnosed cancers in females, breast cancer is about 1/3 of all cases, which is more than double of the second -lung cancer- according to (CDC, 2010). X-ray mammography (including screen-film and full-field digital) has been used to screen for breast cancer since the 1980s. With both mammography modalities and with better treatments, breast cancer mortality rates went down by approximately 24% through 1990s, with both screening and treatment contributing equally to the reduction (Berry et al., 2005).

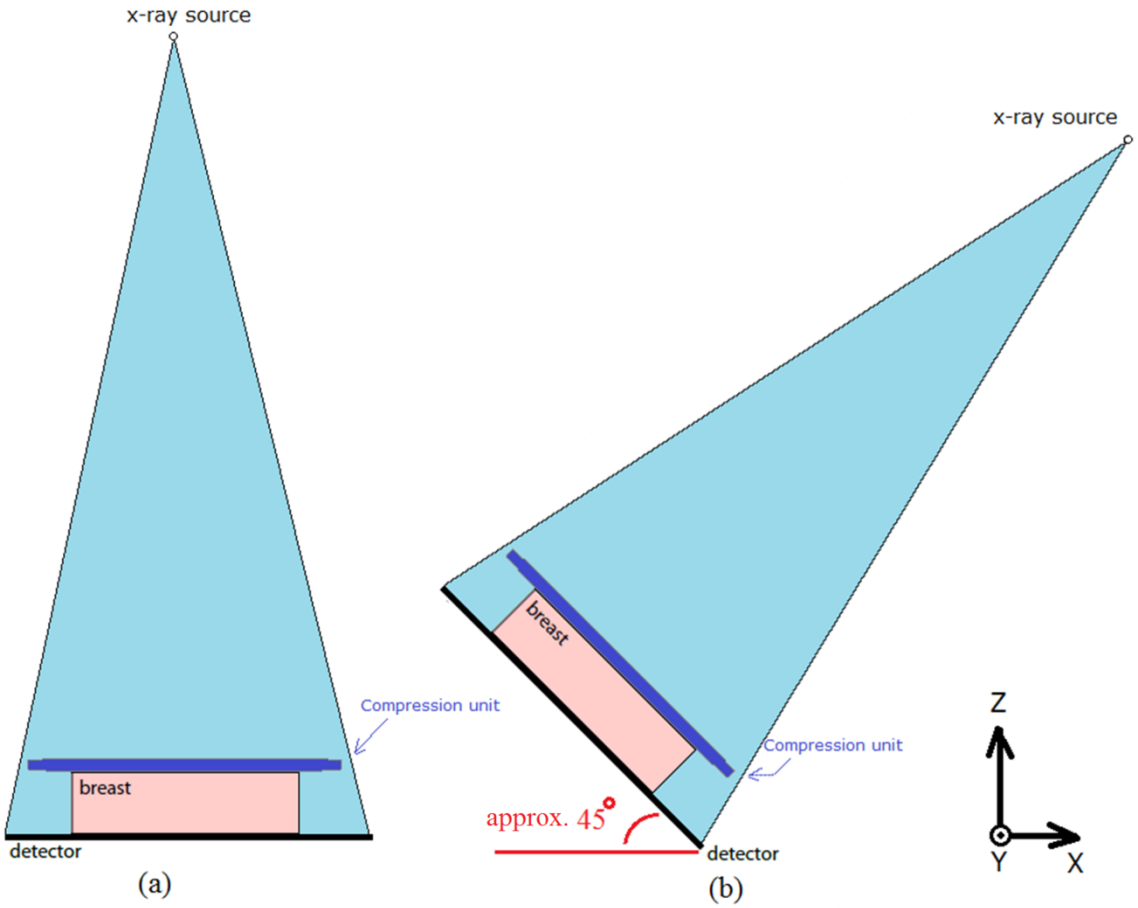

Figure 1.1 (a) Mammography geometry for craniocaudal view, showing fixed x-ray source, phantom, compression unit and detector; (b) Mammography geometry for mediolateral oblique view geometry. The breast appears here as a simple rectangle. This is due to compression which is described below. (c) Coordinate convention. The X-ray tube rotates in the X-Z plane and the detector is in the X-Y plane.

In mammography, one uses low-energy X-rays (usually between 20-30 keV) to take a 2D projection image of a human breast for diagnostic or screening purposes. In screening mammography, both a head-to-foot (craniocaudal, CC) view as seen in Figure 1.1 (a) and an angled side-view (mediolateral oblique, MLO) as seen in Figure 1.1 (b) images of the breast are taken. In the CC view, the breast is oriented so the chest wall is parallel to the X-Z plane and the chest-nipple direction is in Y. Figure 1.2 shows a commercial mammography unit with a patient being imaged in the MLO view.

The X-rays are generated by an X-ray tube by bombarding rotating anode with high energy electrons and generating Bremsstrahlung radiation. The generated X-rays have a broad energy spectrum. With a filter made of some specific material placed at the output of the tube, the X-rays can be filtered to fit into a certain desired energy range. The filtered X-rays (X-ray photons) then interact with human breast by mainly Compton scatter and photoelectric effects. X-ray photons that make it through the breast are then be detected by detector (film or a full field digital detector). Note that during the procedure, the breast is compressed using a dedicated mammography unit. Parallel-plate compression evens out the thickness of breast tissue to increase image quality by reducing the thickness of tissue that X-rays must penetrate, decreasing the amount of scattered radiation (scatter degrades image quality), reducing the required radiation dose, and holding the breast still (preventing motion blur).

Figure 1.2 An X-ray mammography unit showing an MLO view acquisition. Photo is in courtesy of Hologic, Inc.

After image acquisition, breast images are sent to radiologists to look for any suspicious indicators of pathology. Two typical kinds of indicators of pathology are masses and

calcifications. Masses appearing to be spiculated or irregular or showing some architectural distortion usually indicate possible cancer as seen in Figure 1.3 (a). Malignant calcifications usually are clustered and pleomorphic (with different shapes and sizes) as shown in Figure 1.3 (b). Typically, if suspicious indicators are seen after a screening imaging, then a follow-up diagnostic image and, if needed, a follow-up biopsy and pathological exam of excised tissue will be done to determine if the indicator(s) is cancer or not. Although mammography made its success as mentioned in the second paragraph in this section, for dense breasts, mammography's sensitivity can be as low as 48% (Kolb et al., 2002) and the call back rate (i.e. rate of needing to do a follow-up diagnostic image) can be as high as 13% (Smith-Bindman et al., 2005), while breast cancer prevalence is only about 0.5% (Lewin et al., 2001) (Lau, 2012).

Mammography projects a 3D object onto a 2D detector. The 2D image then consists of projections of all the connective tissue, ducts, blood vessels and fat. The image that appears is due mainly to the attenuation of adipose (fat) and parenchymal (glandular) tissues. Essentially a 2D image is a superposition of everything. Typical indicators, clusters of calcifications or masses might lie behind a dense piece of the breast in the 2D images, which might lead to an inability to sort out the abnormality from the normal background. One proposal to solve this problem is to acquire multiple views of the breast from different angles as seen in Figure 1.4, and then implement a tomographic reconstruction of the breast to achieve a 3D view which may reveal pathology indicators easily. That is, the radiologist can view the 3D breast one 2D slice at a time to see structures that may be lost in the 2D mammographic view due to tissue overlap. This is digital breast tomosynthesis (DBT). In this proposal, "Tomo" will be used to mean DBT without any confusion. Figure 1.5 and Figure 1.6 show comparisons between conventional 2D mammography and 3D Tomo reconstructed images. Detail can be found in figure captions.

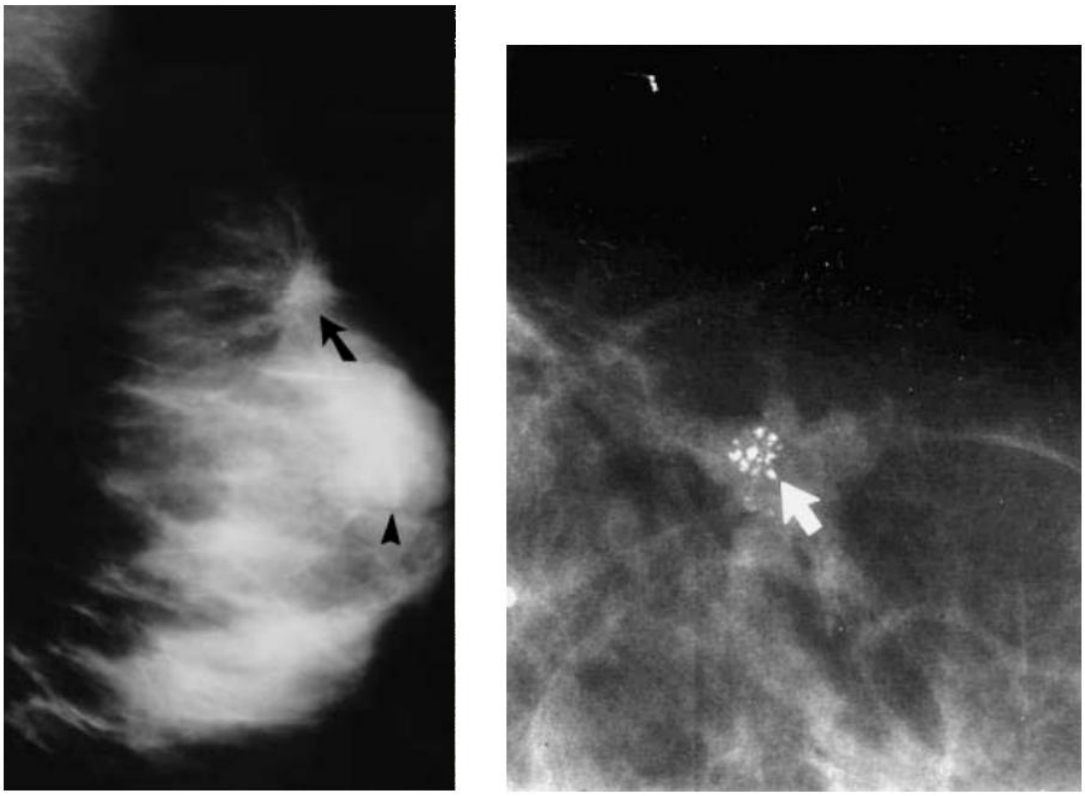

Figure 1.3 Breast cancer lesions. (a) Spiculated mass as black arrow pointed out; (b) Magnified image of clustered, pleomorphic calcifications indicating malignancy. Figures are in courtesy of Radiology Technology 98- Mammography course from Santa Rosa Junior College.

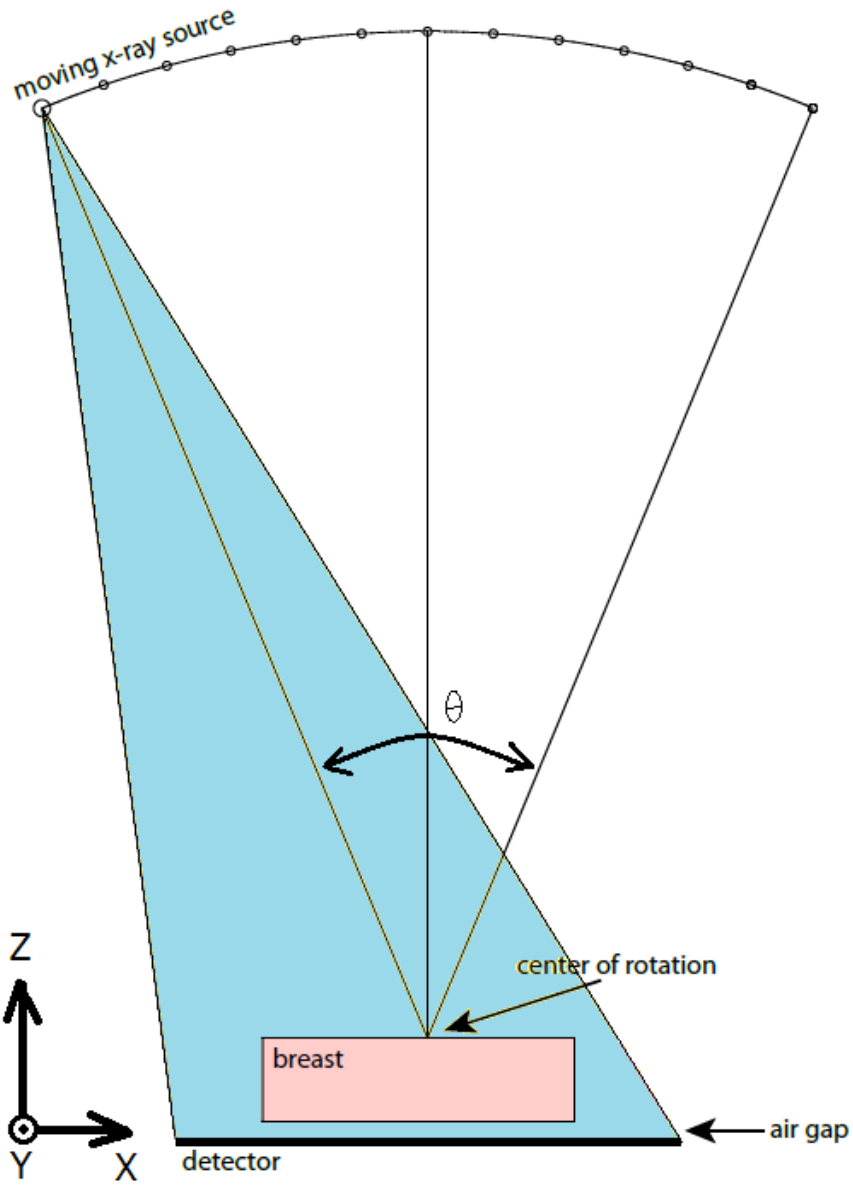

Figure 1.4 Tomosynthesis geometry. An X-Z plane is shown. The moving x-ray source, center of rotation, and air gap are shown. The air gap is the distance from the bottom compression surface to the detector surface. Shaded portion indicates X-ray field from one position. Here θ is the total angle of view. Pictures are drawn to scale, except with fewer angles than the clinical machine we are using. Note that the rays entering the breast from any angle are approximately parallel.

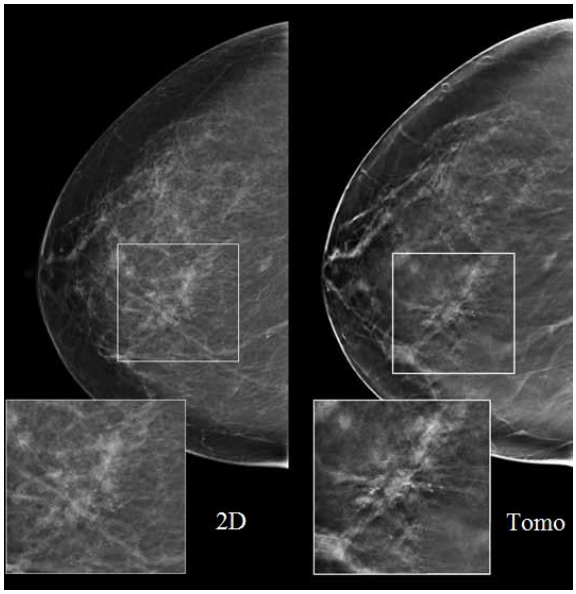

Figure 1.5. The 2D mammogram on the left shows right medial microcalcifications. The tomosynthesis reconstructed slice on the right illustrates the associated architectural distortion only revealed on the CC tomosynthesis image and not on the mammogram. Pictures are in courtesy of HOLOGIC, Inc.

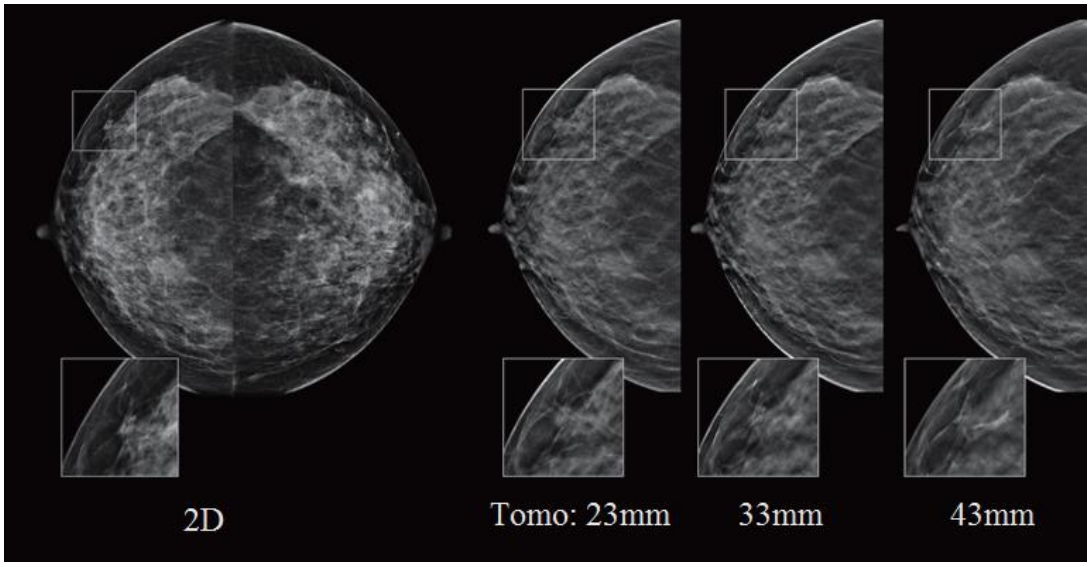

Figure 1.6 Reduced recall rates by tomosynthesis. The 2D mammogram reveals what appears to be a spiculated mass laterally in the right CC view. Tomosynthesis slices at 23, 33 and 43 mm above the breast platform show that this 2D finding was superimposed structures, resolved through the use of tomosynthesis imaging. Pictures are in courtesy of HOLOGIC, Inc.

For practical reasons, it is relatively convenient to modify a conventional mammography unit to a DBT unit. It is relatively cheap to do so. But the angle of view is limited for practical reasons. An experimental DBT unit used by us can be seen in Figure 1.7.

Figure 1.7 The Siemens Mammomat Inspiration DBT unit is shown with the phantom situated for acquisition. The compression paddle sits atop the phantom, and the phantom sits directly atop the carbon-fiber detector cover. The X-ray tube is housed in the top enclosure that rotates.

Other solutions to solve the overlapping problem exist as well: dedicated breast CT (Lindfors et al., 2008), and breast MRI (Houssami and Hayes, 2009) are two. Each modality has its strengths and weaknesses. For instance, breast CT can avoid limited angle reconstruction and provide a 360 degree full angle scan, however, it is expensive to do for screening purpose and also results to be poorer lateral resolution than Tomo. Breast MRI is more sensitive than DBT

and it lacks ionizing radiation, however, its high cost and time consuming scan make it impractical to be more widely used than DBT.

The radiation dose, discussed later, of DBT screening is usually limited to a safe level. For DBT acquisitions, dose is designed to match that of conventional mammography.

So far we have given a very superficial description of DBT. More detail is added as needed in later chapters.

1.3 Basics of SPECT

In Nuclear Medicine, a branch of medical imaging, radioactive material is applied to patients for the purpose of diagnosis and therapeutics. A doctor injects a radiopharmaceutical containing a pharmaceutical tagged with a radionuclide into a patient's body. In SPECT, the radionuclide is an atom in the pharmaceutical that is radioactive and emits γ-rays in the 80-300 keV range. The pharmaceutical goes to an area of interest in the body by biological action. For example, the pharmaceutical may be designed to seek out tumors of a certain kind. It then emits γ-rays which are detected by position sensitive detectors, and an estimate of γ-ray source density is made by a tomographic reconstruction. The reconstruction is a 3-D map of the spatial density of radionuclide. The physician views 2-D slices of the 3-D reconstruction.

In recent decades, two emission tomographic techniques become widely used: single photon emission computer tomography (SPECT) and positron emission tomography (PET). In SPECT, radionuclide emits a single γ-ray in any direction and in one or more energies in the 80–300 keV range. Imaging is accomplished by collimation with lead or other highly absorbing materials. In PET, a positron emitter replaces the γ -emitter of SPECT. The positron is emitted, drifts a few mm, annihilates with an electron to emit two 511 keV γ -rays travelling 180° apart. Since the PET γ -rays fly apart at 180 degrees, no bulky collimation is needed. Instead, timing electronics detect the near simultaneous reception of 511 keV photons, which implies that the annihilation occurred somewhere along the line connecting the detectors. This action of PET is exploited in the instrumentation to derive reconstructions. In this proposal we focus on SPECT.

Applications in SPECT include cardiac imaging, where often the physician looks for a perfusion defect in the heart wall muscle. This shows up as a cold spot (a dip in radionuclide density). (Wernick and Aarsvold, 2004) Another application is in bone imaging (Parker and

Anthony, 2011), where tomography shows up as bright spot of high radionuclide concentration. Other applications include imaging infection, imaging tumors by cell-surface receptors, other forms of cardiac function such as ejection fraction measurements and thyroid fraction imaging. Cardiac studies comprise over half of all SPECT studies. Figure 1.8 shows anatomical drawings of the heart and corresponding reconstructed SPECT slices. For a more comprehensive survey of SPECT applications, see (Wernick and Aarsvold, 2004).

Often the radionuclide used in studies above emits γ-rays at a single energy. For example, Tc99m which emits at 140 keV is used with an appropriate radiopharmaceutical for cardiac (Wernick and Aarsvold, 2004) and bone imaging. In other applications, such as neuroendocrine tomography imaging (Lamberts et al., 1992), the radionuclide, in this case In-111, emits at two energies 171keV and 245 keV.

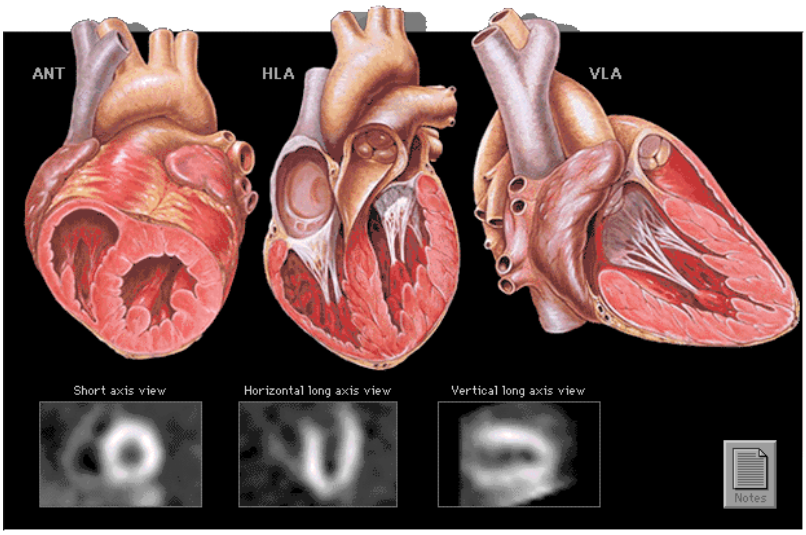

Figure 1.8 Heart image from SPECT. The anatomical drawings are along the top row and the corresponding SPECT image in the bottom row (Yale University School of Medicine, 2004). The bright ring in each image is due to the large uptake in the thick left ventricular muscle.

Planar imaging is the foundation of 3D tomographic techniques just as mammography is the foundation of DBT. Thus, it's necessary to describe in it more detail. In planar imaging, the patient is positioned under a gamma camera. When the radioactive sources (injected by doctor) decay, they emit gamma rays and the position-sensitive detectors of the camera count the gamma

photons which reflect the spatial distribution information of radiopharmaceuticals within the patient.

Assume that the only γ-rays counted are those with directions approximately perpendicular to the detector plane, as illustrated in Figure 1.9. This can be implemented by a collimator that allows only those photons in a certain angular range to pass. A collimator, which we'll discuss in much more detail, is essentially a metal plate with many parallel holes drilled in it. The plate blocks γ-rays, but the holes allow passage to the detector. The γ-rays passing through the collimator will then be converted to visible light using scintillators. Visible light photons positions are then further are decoded by photomultiplier to form 2D projection images. A description of detector technology can be found in (Cherry et al., 2012) and is discussed in more detail in later chapters.

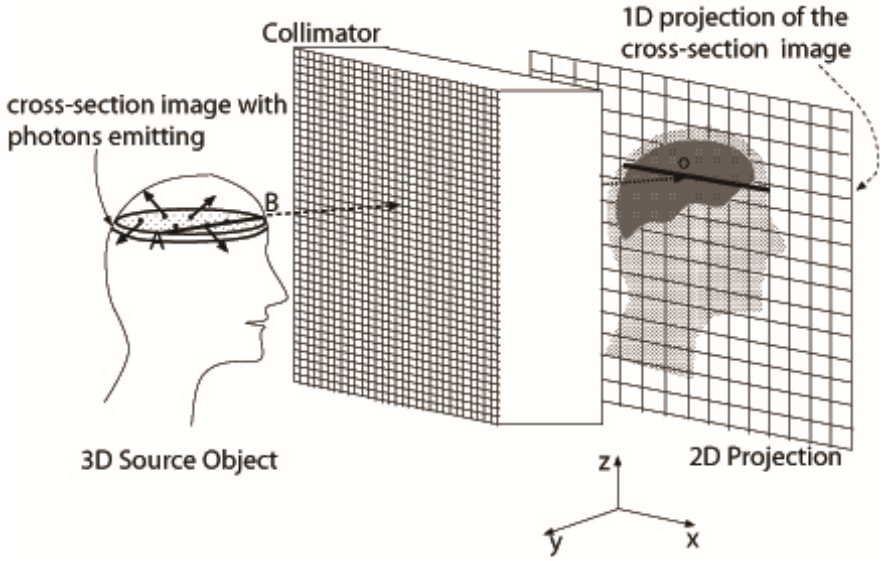

Figure 1.9 A 2D projection from a 3D object. The radionuclide distributes in the 3-D brain and the detector shows a 2-D projection. A particular slice is shown in the brain and its corresponding 1-D projection in the image.

In Figure 1.9, we can see how planar imaging works. Any photon traveling along direction AB will hit the same spot “o” on the detector. As a consequence, all γ-rays emitted along strip AB in the patient contribute to the same detector location “o”, and depth (distance to detector

along y-axis) information is lost. Thus the measurement at “o” is approximately a line integral of the activity along the strip AB. Following this principle, the detected photons from a “slice” cross-section perpendicular to the detector plane (the figure shows one slice of the brain) will form an image along one line, the one-dimensional projection. This is indicated by the dark line in Figure 1.9. A 2D projection is obtained from a 3D object by considering a stack of slices, as displayed in Figure 1.9. Note that the projection is formed at a particular angle about the Z-axis.

Tomography needs projection views at many angles all around the object. By using a mathematical method (reconstruction) to combine them, we can obtain the 3D structure. In SPECT, a gamma camera is designed to rotate up to 360 °around the patient. The camera stops at each equidistant sampling angle to collect gamma photons (projection data). As the projection data are acquired and stored, the computer will assimilate and process the data and do the reconstruction of the original radioactivity distribution of the human body.

This description of SPECT is rather crude and is meant to give the general idea. In later chapters, more detail will be added to this description.

1.4 Introduction to image quality

Improvement of hardware or data processing is done to deliver "better" images. A means of measuring degree of improvement is needed. The image is meant to support a task and degree of improvement should be a scalar measuring task performance.

A simple empirical measure of image quality is CNR, contrast-to-noise ratio. Let's say we had a signal (i.e. a tumor) in a breast image. The breast image comprises scalar values of attenuation coefficient μ_n at each voxel *n*. The tumor may have an elevated value of $\mu = \mu_A$ to background μ_B and noise, characterized simply by a variance σ^2 is present from sources to be described later. The CNR is then

$$CNR = \frac{\mu_A - \mu_B}{\sigma}$$

This measure is crude but useful. In work to be described, we shall make use of these sorts of figures of merits for breast imaging.

A more complicated figure of merit can be explained through a short example: Imagine a nuclear medical planar system (mentioned in section 1.3) used for detecting cancer. The radiopharmaceutical concentrates in the tumor and shows as hot spot. Due to noise in the number of emission γ-rays, images acquired are noisy. One clinical goal is to detect hot spot buried in the noisy image. We could use computer to simulate many acquisition images using a digital phantom that resembles a human body, where the phantom may or may not contain a tumor. As test designers, we know which cases contain a tumor and which do not. To speed up the lesion detection process, we apply a digital observer to decide "tumor present" or "absent" on each image instead of letting a human observer (i.e. radiologist) to make a decision. The decision may be true positive (decide "yes" if tumor present), true negative (decide "no" if tumor absent), false negative or false positive. Given the observer's decisions, the formalisms of detection theory as used in Electrical Engineering can be used to derive some figure of merit for detection performance. In this case, the area under the ROC curve (receiver operating characteristic curve)(Van Trees et al., 2013) is our scalar figure of merit. Given two system, e.g. two SPECT systems with different collimators, we can use this scalar figure of merit to decide which collimator is better.

It turns out that I will use digital forms of observers for SPECT and apply these to optimize the collimator but not the reconstruction. In particular I will use an "ideal observer" (discussed below) applied to raw 2D projection data to optimize the collimator for a novel task in which a signal must be detected and localized. The SPECT work I propose in my dissertation work greatly extends similar earlier work done in my M.S. Thesis (Lu, 2011). In particular, it deals with physically realistic models of SPECT imaging. A strategy of using an ideal observer to optimize the imaging hardware and a separate human emulating observer to optimize the reconstruction images is advocated in (Barrett and Myers, 2003).

Loosely speaking, an ideal observer is one which yields the best possible task performance measure. For example, let's say that in the collimator experiment discussed above, the task was to detect the presence or absence of a signal and let's say that the effect of the collimator was to convolve the true object with a point spread function (psf). The psf sits in noise due to radiation. If the noise were iid Gaussian, then the matched filter is the well known ideal observer. It yields a unique best possible ROC (Receiver Operating Characteristic) curve (Kay, 1998) in terms of area under the curve and other equivalent measures. We will use a different ideal observer, and a

complex task, and a complex collimator and imaging system with Poisson noise (not Gaussian noise). Ideal observers are also used in X-ray imaging and in DBT. These have a simple form due to the modeling of the imaging system as space invariant (blur can be characterized by transfer fraction in Fourier space) and stationary (noise is characterized by a noise power spectrum). Dr. Wei Zhao's group uses a well-known ideal observer for simple detection tasks in DBT for detecting a tumor. They applied it in (Hu et al., 2008) (Hu and Zhao, 2011). That ideal observer is for a task of detecting a signal in the DBT reconstruction. This differs from the one we shall use for detecting/ localizing a signal in SPECT.

1.5 Outline of Thesis

In chapter 2 we talk about details of DBT imaging, focusing in particular on one form of DBT known as CE-DE-DBT. In chapters 3 and 4 I describe the work I have done in scatter correction and joint work I did in reconstruction for CE-DE-DBT. In chapter 5, details of SPECT are given and in chapter 6 I discuss the task-based performance evaluation methodology. In chapter 7 I describe the work I have done in optimizing collimators for SPECT imaging (with In-111 radionuclide) using the method described in chapter 6. In chapter 8, I demonstrate the work I have done in extending the method described in chapter 6 to X-ray transmission tomography. In chapter 9, I summarize the contribution of this thesis.

Chapter 2 A Closer Look at Digital Breast Tomosynthesis

We would like to introduce some relevant concepts in DBT to support work that will appear in Chapters 3 and 4. The material here is meant to help establish definitions and vocabularies that will be used in later chapters. We will discuss about the following topics: imaging procedures and physics involved in mammography/DBT, geometrical setup of the DBT system, imaging equations, reconstruction, use of GPU, an introduction to Contrast Enhanced DBT, and some detail on scatter.

2.1 Imaging chain in mammography/DBT

We will discuss mammography imaging procedures and physics since DBT is essentially a repetition for many angles of these procedures. Much physics involved in mammography can also apply to DBT.

An X-ray tube consists of a vacuum tube, a cathode, a spinning anode and a high voltage circuit as shown in Figure 2.1. After connecting the high voltage between anode and cathode, a stream of electrons will be accelerated shooting from cathode to anode. The accelerated electrons' energy depends on the voltage value. When the electrons reach the anode, typically made of tungsten, a few of them will be deflected and decelerated due to the Bremsstrahlung effect and generate X-rays directed towards the X-ray detector. The typical spectrum of X-rays leaving the tube is shown in Figure 2.2. The broad low component is due to Bremsstrahlung with the highest energy determined by the maximum electron energy. Some bombarding electrons will interact with the K shell electrons of tungsten and trigger narrow-band characteristic radiation peaks. The "amount" of X-rays generated is governed by the product of the tube current and the time that the electron beam is on. This "mAs" (milliamp-second) value is controlled by the operator. A calculated energy spectrum of X-ray photons received is displayed in Figure 2.2. (Beutel et al., 2000).

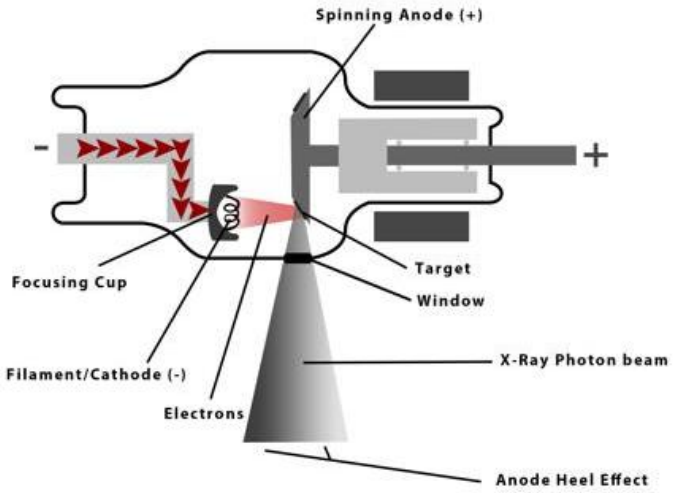

Figure 2.1 X-ray tube diagram. The electrons are accelerated toward the rotating anode. They collide with the anode and generate photons due to Bremsstrahlung and characteristic radiation. The anode geometry is shaped so that the photons travel toward the patient. The intensity across the beam is non-uniform due to the heel effect. The part of the anode that is bombarded at any instant is a thin strip, but due to angulation, this strip appears a small "focal spot" as viewed from the detector. Thus the X-ray, tube can be approximated as a point source of radiation. The figure is in courtesy of radiopaedia.org.

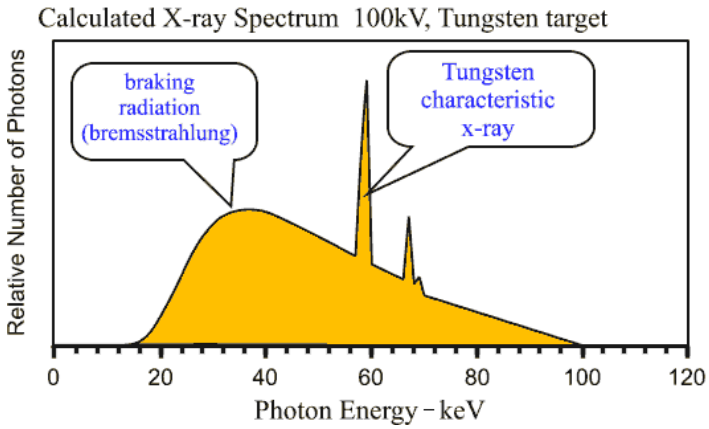

Figure 2.2 Calculated X-ray spectrum generated at 100kVp with tungsten target. The broad spectrum is due to Bremsstrahlung and the narrow peaks due to characteristic radiation.

As one can see, the energy spectrum of the generated X-rays is very wide as shown in Figure 2.2. However, the X-ray spectrum used for X-ray breast imaging is narrower than that seen in Figure 2.2 due to filtration. One example of such a filtered spectrum can be seen in Figure 2.8. The reason for using a narrower band X-ray spectrum is that certain desired energy X-ray photons can create the best contrast for breast imaging and limit the patient dose (Carton et al., 2010). To acquire such an X-ray spectrum, one needs to use a "filter", which is usually a piece of thin metal, e.g. aluminum, to shape the wide X-ray spectrum (Beutel et al., 2000). The filter is placed at the output of the X-ray tube.

As seen in Figure 2.1, the X-rays are emitted, at any one moment, along a stripe from the beveled surface of the anode. From the view of the detector, the stripe looks like a point. This is the "focal spot" of the X-ray tube. Ideally one may expect the spatial intensity of the X-ray emitted from the anode to be uniform. However, due to the complex path lengths of electrons in the anode and the anode geometry, the emitted X-ray beam intensity is non-uniform with angle. This will cause some non-uniformity of the emitting X-ray, which is called the *heel effect* as shown in Figure 2.1 (Beutel et al., 2000).

2.2 Geometrical setup for DBT

A photograph of a particular DBT unit (Siemens Mammomat Inspiration CEDET DBT prototype unit) and its corresponding abstract geometrical diagram are shown in Figure 2.3. One can notice that on top of the breast (here a breast phantom), there is a plastic compression paddle which is used to compress the breast (reasons for compression will be discussed in section 2.3). On the bottom of the breast, there is a carbon fiber detector cover which leaves an air gap between the breast and detector. The distance from the X-ray tube to the center of rotation is 60.85cm. Since the X-ray tube acts as a point source and the cone angle is shallow, the rays hitting the breast are approximately parallel. The detector we use here is an Amorphous Selenium (a-Se) digital flat panel detector. Its dimension is 30.5cm by 24 cm (3584 pixels by 2816 pixels) with square pixel (bin) size equal to 0.085mm. The readout for each pixel of the detector is an electrical pulse converted to a digital value by an analog-to-digital converter. The digitized values are stored in computer memory for later processing and display. It is a direct detector, which directly transfers X-ray photons into an electrical signal. It integrates photons within a certain time slot and the subsequent electrical pulse is sent out. Compared to an indirect detector, which first transfers X-

ray photons into visible light then further converts these to an electrical signal, a direct detector yields better spatial resolution (Zhao, 2007). Note that the complex physics of the X-ray digital detector is out of the scope of my research, but it is a major field of study in its own.

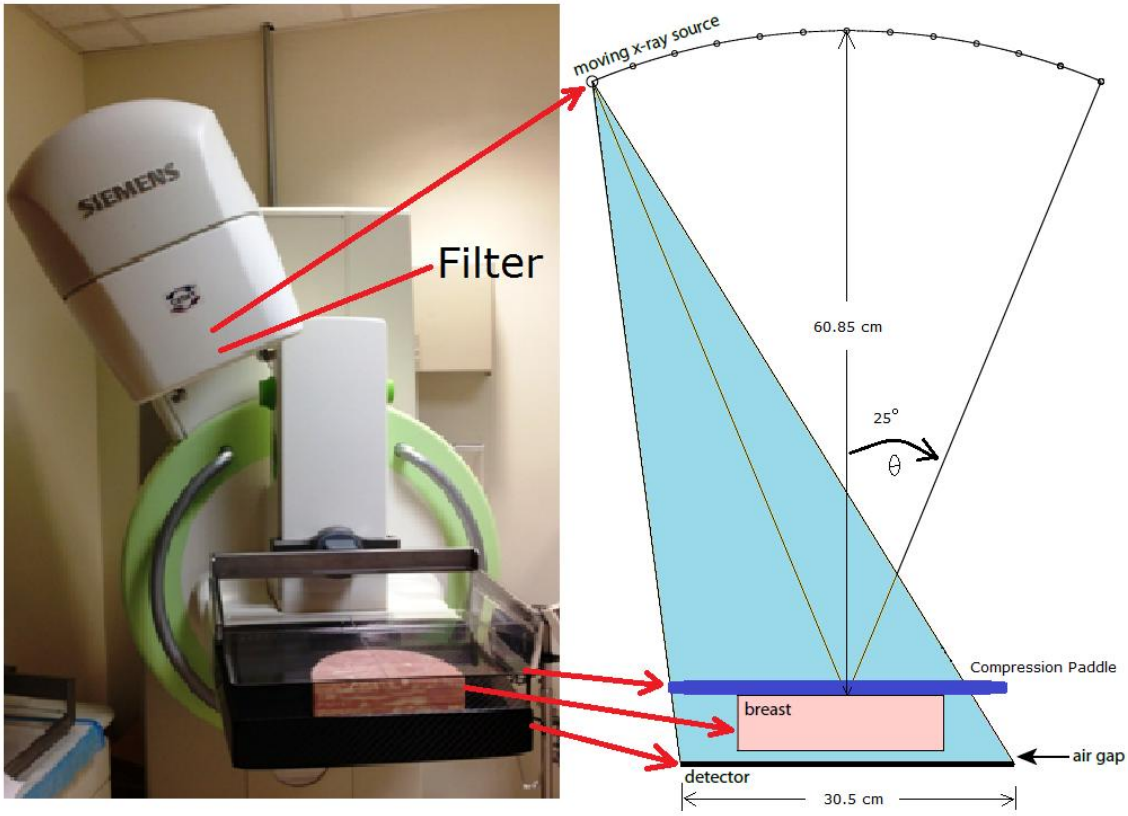

Figure 2.3 An X-ray DBT unit and its geometry. The unit has a phantom (semicircular slab) sitting where the compressed breast would go. One can see the X-ray tube positioned at an angle. The X-ray tube is enclosed within the housing at the top of the unit. The diagram shows the geometry**.**

DBT can be viewed as digital mammography repeated at many angles. As seen in Figure 2.3, an X-ray tube mounted on a moving gantry is used to repeat X-ray imaging at each angle. For the Siemens Mammomat Inspiration DBT unit we are using, it sweeps across from about -25 to +25 degrees. During a DBT imaging acquisition, the tube moves continuously at a constant speed through the angular range. What follows is a typical acquisition sequence. The X-ray tube is fired every 2 degrees so 25 equi-angular exposures are taken. If we divide the total scanning time

(about 20 seconds) evenly into 25 angular segments, then for each divided time slot, we can define a duty cycle. The angular segment for the example here is about 800 ms. Within the angular segment, the X-ray tube fires for about 120 ms. We define this ratio 120ms/800ms as the X-ray tube duty cycle. One can imagine that there is a potential blur in the image just due to the moving X-ray tube. Thus there is blur due to the finite focal spot size and to the motion of the focal spot during X-ray firing.

2.3 Interaction of X-ray photons with breast

The breast is mostly composed of fat and glandular tissue. Other components include areola, nipple and skin. In my study, we are mostly concerned about fat and glandular tissue – the higher proportion of glandular tissue, the denser the breast, which usually leads to a higher risk to have breast cancer.(Kolb et al., 2002, Yankaskas et al., 2001)

Photons that make it through the compression paddle and reach the breast will either directly go through the breast or interact with the breast as shown in Figure 2.4. The two major ways that photons interact with the breast are Compton scatter and the photoelectric effect (Cherry et al., 2012). In the photoelectric effect, the photon is absorbed by the atom and disappears. An electron is generated, but travels only a short distance and does not affect imaging. In Compton scattering, a photon is deflected by an atom and changes direction and loses energy. It generates an electron, which again is of no concern. Thus the photoelectric effect loses photons and the Compton effect can also lose photons. This creates the shadow which is the X-ray image we see. At each point in the breast, there is a different energy dependent attenuation coefficient, i.e. probability to absorb or scatter and X-ray photons. This removal of photons creates the shadow, i.e. contrast in the X-ray image. Some Compton scattered photons may reach the detector and degrade the image quality (Sechopoulos et al., 2007b, Boone and Cooper III, 2000, Wu et al., 2009). The situation is summarized in Figure 2.4.

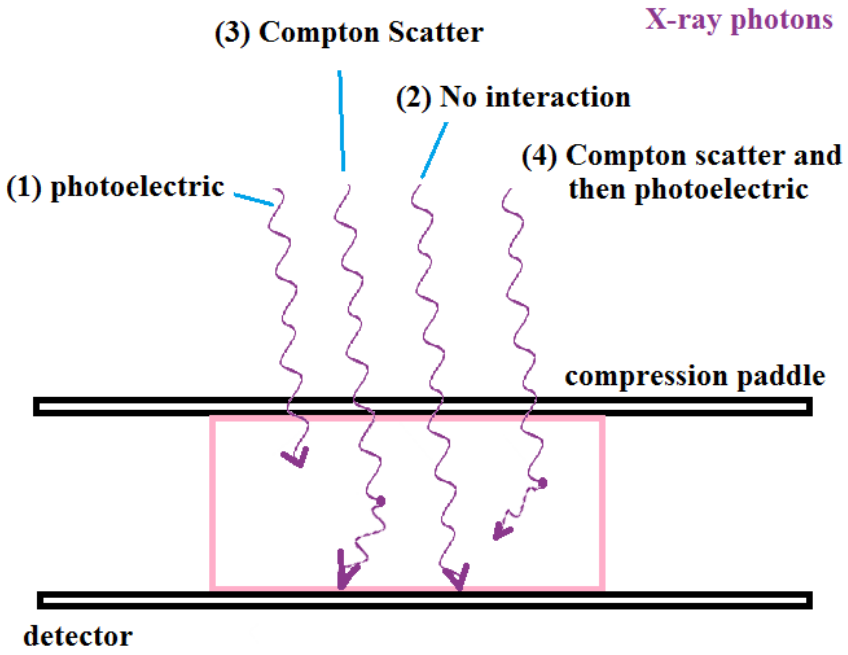

Figure 2.4 Interaction of photons with breast. Four fates are shown (1) Photon passes through breast unaffected (2) Photon is absorbed via the photoelectric effect (3) Photon is Compton scattered then hits detector. (4) Photon is Compton scattered then absorbed by the photoelectric effect.

The Compton scattered photons will have a negative effect on image quality. Since the direction of each scatter event is probabilistic, the sum of the scatter photons will create a low spatial variant "haze" and lower the contrast of the image. For some cases, this spatial variant haze will create some severe "cupping" variations in intensity that will be discussed later in Chapter 3. To remove scatter in mammography, people use an anti-scatter-grid to sort scatter out. A grid is a collimator that rejects scattered photons whose directions do not correspond to a path from the X-ray point source to the detector (Sechopoulos, 2013). However, for DBT, a grid is not used since a grid must move with the tube which is impractical, so scatter is still a problem.

We can divide the four kinds of photons in Figure 2.4 into three groups. One would be attenuated photons corresponding to rays 1 and 4. These do not hit the detector. Another one would be primary photons corresponding to ray 2 which are the photons that do not interact with

the breast and hit the detector. The third group corresponds to ray 3. These are Compton scattered photons that do hit the detector. The primary photons follow a pencil beam Beer's law:

$$I = be^{-\mu x} \tag{2.1}$$

Equation (2.1) applies to a uniform attenuator where the path length of the beam is x. The b is the sum of total photons of a beam before the breast, x is the distance a beam travels inside the breast, I is the primary photons flux. The μ is the attenuation coefficient, and has units of cm^{-1}. It can be shown that $\mu = \mu_{scatter} + \mu_{photoelectric}$. This description is quite basic. In fact, μ depends on energy and varies with spatial position. The path length x actually depends on the focal spot position and detector element position. The simple Beers law above is replaced by a more complete description and equations below.

The breast compression lowers scatter, minimizes tissue overlap in the projection image, and keeps the breast from moving. It is painful so the amount and time of compression should be minimized (Broeders et al., 2015).

We will also mention dose. Dose is a complex topic (Samei and Saunders, 2011, Sechopoulos et al., 2007a, Boone et al., 2001), but qualitatively, it is defined as the energy deposited in the tissue by the beam via the various physics effects we discussed. A high dose can do biological damage to cells and cause cancer. (So detecting cancer can cause cancer!) Dose is measured in units of milliGrays (mGy). Dosimetry is out of the scope of this thesis. We consider it indirectly in our experimental acquisitions by using an mAs which is consistent with recommended dose constraints.

2.4 DBT imaging equations

If we take an X-ray point source image without the breast, we call it a *blank scan*. A blank scan is a 2-D image is that subject to inverse square shading, the heel effect and some other effects that make the full detector field intensity non-uniform. One other effect is foreshortening. A beam arriving at θ sees a pixel at angle θ whose projected area is reduced by $\cos\theta$. Therefore the intensity is similarly reduced. We define the 2-D blank scan image using a vector

$\mathbf{b}^{\theta}$, θ is the X-ray tube's angle and $\mathbf{I}$ is a digitized signal out of the detector. (Note we use "vector" in the sense that $\mathbf{b}^{\theta}$ is a 2-D array of detector readings, which when lexicographically ordered, form a list of numbers comprising a vector.) So if we did a DBT study with no breast we would collect images $\mathbf{b}^{\theta}$.

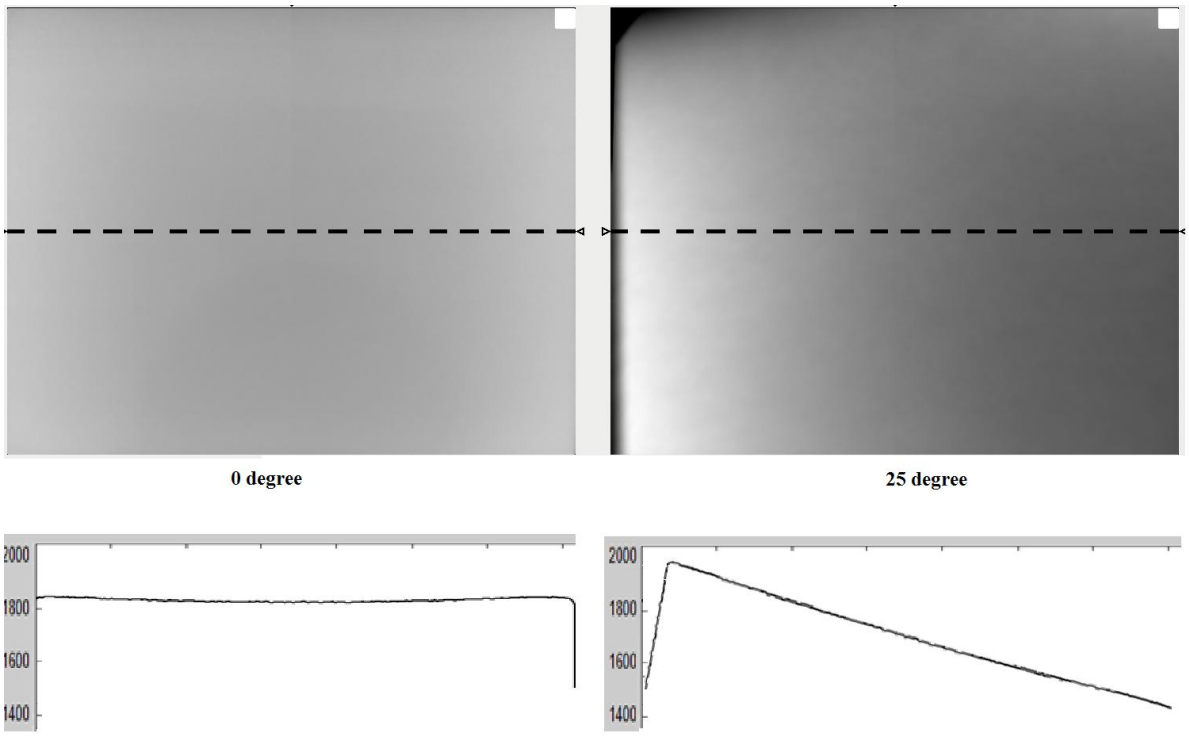

Figure 2.5 Blank scans at two angles. Gray scale images are two blank scan at 0 and 25 degrees. Profiles across the grayscale image (located at the dotted line) sit below grayscale images. The images were recorded from our scanner.

Beer's law applies to the various rays hitting the breast. Let *m* index pixels on a detector, and *n* index the object (breast) voxels so μ_n is the attenuation coefficient of the *n*th voxel. Let A_{mn}^{θ} be the element of matrix $\mathbf{A}$ that indicates path length through the voxel *n* to pixel *m* at angle θ. Then Beer's law is extended to

$$\left[\mathbf{I}^{\theta}\right]_m = \left[\mathbf{b}^{\theta}\right]_m e^{-\sum_n A_{mn}^{\theta} \cdot \mu_n} \tag{2.2}$$

where $[\mathbf{I}]_m$ is the *m*th component (*m*th detector element reading) of $\mathbf{I}$. In matrix-vector shorthand this becomes $\mathbf{I}^\theta = \mathbf{b}^\theta e^{-[\mathbf{A}\boldsymbol{\mu}]^\theta}$. We call $\mathbf{A}$ the system matrix and it captures the geometry of the system. The $\mathbf{I}^\theta$ comprise the collection of 2D images (one per value of θ) obtained in a DBT study. However equation (2.2) is still approximate.

As mentioned in section 2.3, scattered photon do not follow a simple path. Let $\left[\mathbf{s}^\theta\right]_m$ be the scatter contribution to pixel *m* with the X-ray tube at θ. Then we can extend the imaging equation to

$$\left[\mathbf{I}^\theta\right]_m = \left[\mathbf{b}^\theta\right]_m e^{-\sum_n A_{mn}^\theta \cdot \mu_n} + \left[\mathbf{s}^\theta\right]_m \tag{2.3}$$

where $\mathbf{s}^\theta$ is a vector whose elements are $\left[\mathbf{s}^\theta\right]_m$.

So far, we have not spoken of noise. The number of photons in the X-ray beam has an irreducible component of Poisson noise as does any electromagnetic beam. Furthermore the detector has electronic and readout noise. We lump those together to get a noise component

$$\left[\mathbf{I}^\theta\right]_m = \left[\mathbf{b}^\theta\right]_m e^{-\sum_n A_{mn}^\theta \cdot \mu_n} + \left[\mathbf{s}^\theta\right]_m + \left[\mathbf{n}^\theta\right]_m \tag{2.4}$$

where $\mathbf{n}^\theta$ is a vector of noise contributions to the detector from the acquisition at θ. It has components $\left[\mathbf{n}^\theta\right]_m$. This imaging model is a big simplification in several ways. It does not take into account blur mechanisms or noise propagation mechanisms from one stage to another in the imaging procedures (Zhao and Zhao, 2005, Zhao and Zhao, 2008). Also $\boldsymbol{\mu}$ is dependent on energy, and so this simplification ignores polychromatic effects. Nevertheless, it will be a useful starting point for the work discussed in Chapters 3 and 4.

So far the imaging model has described the formation of projection (line integral) data. To get a 3D estimate of $\boldsymbol{\mu}$, we need to tomographically reconstruct the projection data.

2.5 Reconstruction

We describe 2 well-known algorithms FBP and SART, and point out their advantages. We also describe the general nature of a third type of reconstruction based on statistical estimation. This type of reconstruction will appear in Chapter 4 in the context of an OSTR (Ordered Subsets Transmission Reconstruction) algorithm.

In the previous section, we wrote the imaging equation (2.4). We can first transform the imaging equation to a linear matrix-vector equation as follows: First, we assume that a scatter correction method (the subject of Chapter 3) has removed scatter so the scatter term in (2.4) disappears. Next we divide the equation (left term and right term) by $\mathbf{b}^{\theta}$, then take a log transformation and set it equal to $\mathbf{g}^{\theta}$ to obtain $\mathbf{g}^{\theta} = [\mathbf{A}\boldsymbol{\mu}]^{\theta} + \tilde{\mathbf{n}}^{\theta}$. Note that the $\tilde{\mathbf{n}}^{\theta}$ is the noise term $\mathbf{n}^{\theta}$ propagated through the division-log transformation, so noise is very complicated. We will know $\mathbf{b}^{\theta}$ through measurement, so it can be removed in the transformed equation. For ease of description, we drop the $\tilde{\mathbf{n}}^{\theta}$ term. Finally, we get

$$\mathbf{g} = \mathbf{A}\boldsymbol{\mu} \tag{2.5}$$

where for convenience we have temporarily dropped the θ superscript. It is understood that (2.5) is an inconsistent system due to noise (Even without noise $\mathbf{g} = \mathbf{A}\boldsymbol{\mu}$ is not solvable since $\mathbf{A}$ has a huge null space due to limited angle projection.). Knowing $\mathbf{g}$ and $\mathbf{A}$ our task is to get $\hat{\boldsymbol{\mu}}$, an estimate of the "true" $\boldsymbol{\mu}$. (The true $\boldsymbol{\mu}$ is spatially continuous, and $\hat{\boldsymbol{\mu}}$ is a discrete approximation)

First we describe a version of the well known filtered backprojection algorithm (FBP) adapted to DBT (Mertelmeier et al., 2006). We can use this version of FBP to reconstruct slices in the X-Y plane. We can reconstruct each X-Y plane (i.e. for each Z value) then stack them together to get the 3D FBP reconstruction.

Our DBT hardware uses a stationary geometry in which the detector is motionless. (This is the most common DBT geometry.) The cone of rays emitted from the X-ray tube has such a shallow cone angle that we can consider the rays as parallel as they enter the breast. Given a 2D stationary-detector geometry, a projection $\mathbf{g}$ at 2 angles looks as depicted in Figure 2.6

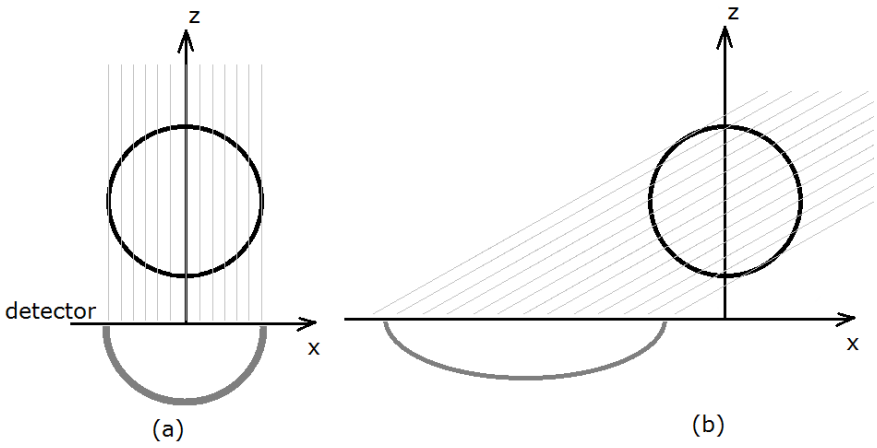

Figure 2.6 Projections $\mathbf{g = A\mu}$ at two angles onto a stationary detector (a) Projection at 0 degree (b) Projection at a large angle. The projection at the large angle is stretched relative to the one at 0 degrees, and also has a lower intensity due to $\cos\theta$ forshortening.

Since the detector is stationary, its projection at a steep angle is wider and of lower intensity than that at 0 degrees due to simple geometry and radiometry. In traditional tomography such as CT, the detector rotates with the source so that all rays at any angle are perpendicular to the detector (an isocentric geometry). Also, projections in traditional tomography are taken over 2 π angles, but in DBT, we span a limited angle range $\left[-\theta_{\max}, \theta_{\max,}\right]$ degrees. The FBP for the traditional case is well known and is found in textbooks, but for the stationary geometry it becomes more complicated.

We will use a continuous representation to describe the DBT-FBP algorithm. Let $g(x,\theta)$ be the projection at θ. Since we are describing a 2D version of the algorithm, we can take the detector to be 1-D and coincident with the x-axis of our DBT geometry. Note also that we have reintroduced θ as an argument to g. We give the steps of the reconstruction without derivations, but see (Mertelmeier et al., 2006) :

(1) Take the Fourier transformation of each projection $G(w,\theta) = FT\{g(x,\theta)\}$

(2) Filter each projection by

$$H(\omega,\theta) = 2\tan(2\alpha)\cdot\frac{|\omega_x|}{\Gamma\cos\theta} \qquad |\omega_x| < \Gamma \tag{2.6}$$

where $\alpha = 2\theta_{\max}$ is the total angular range (Here we used a symmetric angular schedule without loss of generality). The factor Γ is the Nyquist limit of the detector determined by the density of detector elements. The "ramp filter" $|\omega_x|$ is well known in conventional FBP, but for the stationary version, it gets modified to $H(\omega,\theta)$. For noise suppression, one could also multiply $H(\omega,\theta)$ by an apodization filter $A(\omega,\theta)$ but we skip this formalism here.

(3) The filtered projections are $G(\omega,\theta)H(\omega,\theta)$. Take an inverse Fourier transform to obtain filtered projections in the spatial domain $q(x,\theta) = FT^{-1}\{G(\omega,\theta)H(\omega,\theta)\}$.

(4) Backproject and add each $q(x,\theta)$ along the rays from which they came as seen in Figure 2.7.

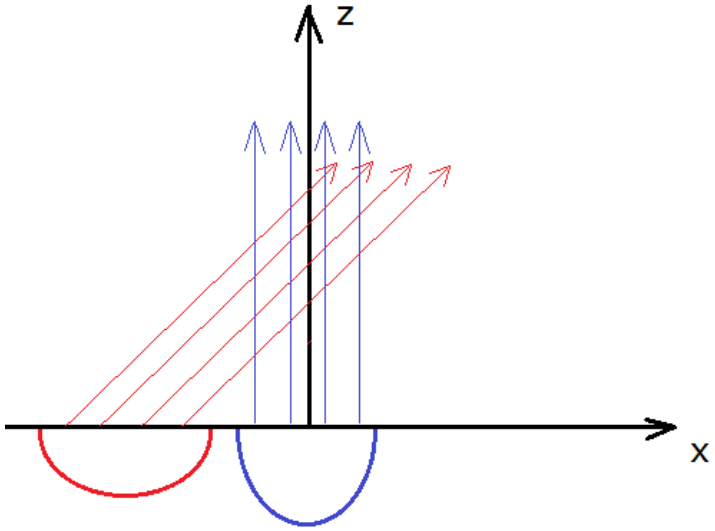

Figure 2.7 Backprojection shown for 2 angles. The backprojection is over the limited angular range $-\theta_{\max}$ to $\theta_{\max}$. Here, a backprojection is shown for two angles, 0 degrees and $\theta_{\max}$.

In practice we discretize the FBP and use FFTs instead of Fourier transformations. FBP is very fast requiring only one backprojection. Theoretically FBP is justified only if we have acquisitions at all angles, but the limited angular acquisition results in blurring in the Z direction for $\hat{\mu}$.

We now switch to a discrete representation to describe SART (Simultaneous Algebraic Reconstruction Technique). (The FBP algorithm was described in continuous space but in

implementation uses this same discretization.) We revert to the matrix-vector notation without the θ superscript. SART attempts to solve (2.5), the undetermined (due to limited angle) and inconsistent system $\mathbf{g} = \mathbf{A}\boldsymbol{\mu}$. The SART algorithm is one of a class of iterative algorithms that generates estimates $\hat{\boldsymbol{\mu}}^k$ at iteration *k*. We represent each element of $\hat{\boldsymbol{\mu}}^k$ by a rectanguloid voxel of value $\hat{\mu}_n^k, n = 1,...,N$ that is approximately 10 times higher (in Z) than it is wide. The discretized space of $\hat{\boldsymbol{\mu}}$ comprises 2000 by 1000 by 30 voxels. The dimension in X Y Z of a typical voxel is 85 μm by 85 μm by 1000 μm. Due to the limited angle acquisition, resolution in the Z direction is far worse than in the X and Y directions, hence the rectanguloid voxels. Vector $\mathbf{g}$ comprises elements $g_m, m = 1,...,M$, one per detector element. Therefore matrix $\mathbf{A}$ is huge: M by N, where M is the number of detector elements, 3584 x 2816, and N=2000 x 1000 x 30. We do not explicitly represent $\mathbf{A}$, but compute matrix vector products on the fly.

The iterative algorithms including SART will, at each iteration, require projection $\mathbf{A}\hat{\boldsymbol{\mu}}^k = \hat{\mathbf{g}}^k$ and backprojection $\mathbf{A}^T\mathbf{g}^k$ operations. (Here T means transpose). Backprojection assigns $\hat{g}_m^k$ back along the ray from which it came, so that addition to voxel *n* is weighted by A_{mn}, the chord length of ray *m* through voxel *n*.

In the applied math literature, there are many iterative algorithms for solving linear systems. SART is one version useful in tomographic imaging. SART is slow and has been modified to a speedier "ordered subset" version. For a background in SART, ordered subset SART and related approaches for reconstruction please see (Wang and Jiang, 2004).

The version of SART we will use in experiments is an ordered subset SART. It is written as an iterative update here. Note in this SART description, the index *m* covers all detector pixels and all angles, so θ is not needed. The reason for this is that it makes it easier to describe and notate the ordered angular subset part of SART.

$$\hat{\mu}_n^{(k,l+1)} = \hat{\mu}_n^{(k,l1)} + \frac{\lambda}{\sum_{m' \in S_k} a_{m'n}} \sum_{m \in S_l} \left(a_{mn} \frac{g_m - \sum_{n=1}^{N} a_{mn} \hat{\mu}_n^{(k,l)}}{\sum_{n=1}^{N} a_{mn}} \right) \tag{2.7}$$

There are two iteration loops, the outer one indexed by *k* and inner one indexed by *l*. The quantity a_{mn} is a matrix element of $\mathbf{A}$. The quantity S_l indicates a limited subset of all angles, and index *l* is updated after each subset has been processed. When all subsets are complete, *k* is

incremented. The factor λ is an acceleration factor typically $0<\lambda<2$ (we use λ=0.2). Note that the quantity $\sum_{n=1}^{N} a_{mn}\hat{\mu}_n^{(k,l)}$ is a projection and that

$$\sum_{m\in S_l}\left(a_{mn}\frac{g_m-\sum_{n=1}^{N}a_{mn}\hat{\mu}_n^{(k,l)}}{\sum_{n=1}^{N}a_{mn}}\right) \tag{2.8}$$

is a backprojection. We have used SART in our lab (Zhou, 2007) in related DBT work. We run SART for $k_{\max}$ = 4 iterations to get a reasonable reconstruction.

Having introduced FBP and SART algorithms, we briefly talk about the quality of reconstructions using them: FBP tends to give an edge enhanced image, which is ultimately due to the ramp filter. For the very same reason, FBP will also suppress the low frequency component of the reconstructed image. Instead, SART can give much nicer reconstruction results, e.g. minimal edge enhancement, low frequency maintenance. In Chapter 4, we will use these two algorithms and do some quantitative comparisons.

Yet another approach is statistical reconstruction. We give a brief motivation here but use it in Sec.4. Historically, a form of transmission tomography was used in the 1990' (Fessler, 2000) in which an external radioactive source replaced the X-ray tube. This form of transmission tomography was useful in PET (Positron Emission Tomography) attenuation correction. In this case, eq.(2.4) holds accurately and $\mathbf{n}$ is independently Poisson at each detector bin m. Polyenergetic effects of X-ray tube emission are gone since we replace the tube by a narrow-band radionuclide emitter. The received signal $\mathbf{I}$ is an integer since we can use photon counting. For convenience we rewrite the equation dropping the θ superscript. By dropping θ, we can presume that m runs over all detector positions and angles, thus leading to less complicated notation:

$$\mathbf{I}_m=\mathbf{b}_m e^{-[\mathbf{A}\boldsymbol{\mu}]_m}+\mathbf{s}_m+\mathbf{n}_m \tag{2.9}$$

where $[\mathbf{A}\boldsymbol{\mu}]_m$ is the mth component of the matrix-vector product $\mathbf{A}\boldsymbol{\mu}$. It can be shown that $\mathbf{n}_m$ is Poisson and is in fact independently Poisson at each m. If that is true then

$$\bar{\mathbf{I}}_m=\mathbf{b}_m e^{-[\mathbf{A}\boldsymbol{\mu}]_m}+\mathbf{s}_m \tag{2.10}$$

is the mean measurement at m. Since n_m is Poisson, we can say

$$p(\mathbf{I}_m \mid \boldsymbol{\mu}) = \frac{e^{-\bar{\mathbf{I}}_m} \bar{\mathbf{I}}_m^{\mathbf{I}_m}}{\mathbf{I}_m!} \tag{2.11}$$

where $\mathbf{I}_m$ is an integer. Since the $\mathbf{n}_m$ are independently Poisson

$$p(\mathbf{I} \mid \boldsymbol{\mu}) = \prod_m \frac{e^{-\bar{\mathbf{I}}_m} \bar{\mathbf{I}}_m^{\mathbf{I}_m}}{\mathbf{I}_m!} \tag{2.12}$$

Note on the right side of eq.(2.12) that $\boldsymbol{\mu}$ is hidden in the term $\bar{\mathbf{I}}_m$.

If we know $\boldsymbol{\mu}$, then eq.(2.12) is the probability of receiving a particular measurement $\mathbf{I}$. On the other hand, if we observe $\mathbf{I}$ but desire to estimate $\boldsymbol{\mu}$ then, with a slight abuse of notation, eq.(2.12) is a likelihood for $\boldsymbol{\mu}$. In this case, we can get an estimate for $\boldsymbol{\mu}$, $\hat{\boldsymbol{\mu}}$, using maximum likelihood (ML) methods. These maximum likelihood methods can be extended to MAP (maximum a posteriori) methods. The problem then reduces to an *algorithm* to find the ML or MAP solution for $\hat{\boldsymbol{\mu}}$, but that is the subject for Sec.4.

Because this approach worked well in the PET context where $\mathbf{n}$ was actually Poisson, it was applied to DBT (Wu et al., 2004, Zhang et al., 2006) using an EM algorithm for ML, even though in DBT $\mathbf{n}$ has non-Poisson detector and electronic noise component in addition to photon noise. In our work, we will apply statistical methods to a particular form of DBT, contrast-enhanced dual-energy DBT, and use an algorithm OSTR (Ordered Subsets Transmission Reconstruction) to solve a MAP version of the problem.

2.6 GPU acceleration of reconstruction algorithms

Parallel computing has become the standard for high-performance computing. Though single-core processors nearly doubled their performance every 18 months over the past thirty years (Chien and Karamcheti, 2013), excessive power dissipation at GHz clock rates and the structure of problems in graphics made developers shift their interest to parallel computing architectures. In the medical imaging community, intensive computing needs due to an exploding amount of data in 3-D or even 4-D in imaging and improvement in spatial resolution still can hardly be fulfilled. Fortunately, graphics processing unit (GPU), which were originally designed for accelerating the production of computer graphics, turned out to be very useful in medical

imaging, especially in the computationally demanding projection and backprojection operations that characterize tomographic imaging modalities (Pratx and Xing, 2011). The same reasoning applies in GPU acceleration in fluid dynamics, differential equations and many other application areas. It includes high memory bandwidth and computation throughput, floating-point arithmetic support and a low price per unit of computation.

In my research dealing with DBT, the ultra-high spatial resolution of the projection image (3584 x 2816 pixel^2) makes a 3D reconstruction computationally extremely time consuming if one uses a regular C programmed package running on a CPU, e.g. it takes more than 6 hours to finish an iterative reconstruction (e.g. SART) for a 2.4GHz CPU. That is because for a given angle, the number of rays equals to number of detector elements. Each ray requires a projection and backprojection at each iteration, so the computational demand is extreme.

Thus, I decided to learn and write GPU programs to help accelerate the reconstruction process. Different from regular C language, a single-instruction single-processing, CUDA (Compute Unified Device Architecture) C/C++ developed by NVIDIA is a single-instruction multi-processing language which programs the GPU to compute. GPU coding disobeys the intuition of common programming logic, which makes it very hard to start and learn. However, the speedup is so high that I was compelled to implement a GPU based reconstruction. After nearly half a year's learning and work, I programmed a CUDA Tomosynthesis reconstruction package: *MIPL-TomoG*. With the aid of GPU, CPU based 6 hours SART reconstruction has been shortened to 5 minutes (a speedup of nearly 100 times!).

The basic idea of GPU computing is parallel computing , it splits one giant computing work into ~100 *independent* threads and computes them simultaneously, then combines the results in the end. One can notice that in the reconstruction algorithm, projection/backprojection needs to be done for each voxel/pixel. Those voxels/pixels are essentially independent. Thus, reconstruction becomes a perfect application of GPU.

Note that my work in GPU programming is not intellectually innovative in that others have written GPU based projection and backprojection software. It is rather an infrastructure building. One can find more about GPU applications in medical imaging in (Pratx and Xing, 2011). Without the speedup due to my GPU implementation for reconstruction, none of the tomo projects I describe in Chapters 3 and 4 would have been possible, so the learning and programming effort were worth it.

2.7 Dual energy breast imaging

Sometimes a cancerous tumor in the breast might be present, however, the contrast of the mass might not be large enough to be seen by the radiologist. Similarly, a mass might show up but may be difficult to be classified as cancer or non cancer for both DBT and mammography. One technique proposed to help is iodine contrast enhanced imaging. Though the exact clinical use of this form of contrast imaging is still being developed (Dromain et al., 2012, Puong et al., 2007), it appears that it may help in increasing specificity (ruling out suspicious area of cancer) or staging the disease when present (Jong et al., 2003). Since it is an imaging modality, it could be done before an invasive biopsy is needed. In iodine contrast enhanced imaging, a contrast agent containing an iodine compound is injected into the patient's bloodstream. Iodine has a high attenuation coefficient for X-rays at relatively high energy, which helps the tumors show up in the following way: The injected iodine tends to concentrate in the tumor region because, due to the increase of metabolism of cancer cells, they tend to grow their own dense network of blood vessels, a capillary network much denser than that in the surrounding tissues. This effect is known as angiogenesis (Dromain et al., 2006). The dense capillary network provides the additional nutrients for the increased metabolism. If iodine is in the capillary system, the enhanced concentration of iodine in the dense capillaries will be associated with tumor (Lu et al., 2013). Furthermore, there's another dimension - the temporal dynamics of iodine contrast, which means how the iodinated blood stream washes in and washes out. The nature of cancer is that the blood vessel is "leaky", and it is thought that iodine concentration washes out more quickly than in normal tissue (Dromain et al., 2006). So if the suspected iodinated region is viewed over several time points, a short washout time may indicate cancer. This effect is still being studied and the use of short washout times as a cancer indicator is still under study (Dromain et al., 2012, Jong et al., 2003).

One reason that contrast dynamics with iodine may be useful is that it is already used in MR (magnetic resonance) breast imaging (Bellazzini et al., 2015, Silverman et al., 2009, Paul et al., 2007). For MR, instead of iodine, another enhancement agent is used- Gadolinium. It's been well-known and used frequently. Research has been done to show that in MR the washout is faster for cancer tumor. So one may ask why not just use contrast enhanced MR? The answer is that it is a quite expensive procedure to do whereas the iodine contrast imaging in X-ray is far cheaper and easier to do. Yet another advantage is that the contrast imaging with iodine in other

X-ray imaging modalities (i.e. CT) is rather common, and it's nothing new. DBT and conventional mammography are just new applications for iodine contrast imaging. What we need is just a conventional mammography unit or a DBT unit, plus a certain iodine injection unit which is rather common. Indeed, we've already done some patient studies here at Stony Brook using the X-ray contrast imaging procedure.

We will describe the basics of CE (contrast enhanced) imaging with the aid of Figure 2.8. The figure shows the low-energy (LE) X-ray spectrum and high-energy (HE) spectrum after filtration and just before the beam hits the breast. The LE spectrum would be the one used in conventional breast imaging. Let us first consider the case of no contrast agent. We'll first show that a form of weighted subtraction of the HE and LE projection images yields a nearly featureless constant projection image devoid of any structure.

A conventional mammogram is due to the attenuation difference between adipose and parenchymal tissue. Tumor cells, such as an invasive ductal carcinoma (IDC), have about the same attenuation as parenchymal cells, and may show up due to their spatial pattern more than due to a vast difference in attenuation. In Figure 2.8, the adipose, fibroglandular and IDC attenuation vs. energy curves are shown, with IDC and fibroglandular essentially overlapped (Note that "parenchymal" and "fibroglandular" are used interchangeably here). On the log scale in Figure 2.8, the attenuation difference of fibroglandular and adipose looks small, but it is this difference that accounts for conventional mammographic images acquired at LE.

Now we will show how weighted subtraction "erases" the conventional projection image. We take as our imaging model eq.(2.5) and modify it for our purposes. Consider the LE case, and consider a monoenergetic approximation where the LE value is the mean energy of the LE spectrum 18 keV, and for HE it is 37 keV. Also consider the idealized case of a breast of thickness *t* composed only of fibroglandular (parenchymal) and adipose tissue. A given ray will traverse the breast and see a fraction α of fibroglandular tissue and hence a fraction $(1-\alpha)$ of adipose tissue. Of course, α depends on the ray and is space dependent $\alpha = \alpha_m$, but for convenience we suppress this dependence. Finally we use subscripts LE and HE to denote quantities as belonging to low or high energy.

Putting all these together, eq.(2.5) becomes for high and low energies

$$g_{HE} = t\alpha\mu_{HE}^{G} + t(1-\alpha)\mu_{HE}^{A} \tag{2.13}$$

$$g_{LE} = t\alpha\mu_{LE}^{G} + t(1-\alpha)\mu_{LE}^{A} \tag{2.14}$$

where superscripts *G* and *A* denote fibroglandular and adipose. We can rewrite (2.13) and (2.14) as

$$g_{HE} = \alpha(t\mu_{HE}^{G} - t\mu_{HE}^{A}) + t\mu_{HE}^{A} \tag{2.15}$$

$$g_{LE} = \alpha(t\mu_{LE}^{G} - t\mu_{LE}^{A}) + t\mu_{LE}^{A} \tag{2.16}$$

The weighted subtraction scheme with weight *w* forms a difference image Δg as follows

$$\Delta g = g_{HE} - w \cdot g_{LE} \tag{2.17}$$

$$\begin{aligned} &= \alpha(t\mu_{HE}^{G} - t\mu_{HE}^{A} - wt\mu_{LE}^{G} + wt\mu_{LE}^{A}) \\ &\quad + t\mu_{HE}^{A} - wt\mu_{LE}^{A} \end{aligned} \tag{2.18}$$

The trick is to determine *w* so as to cancel the term in ( ) in eq.(2.18), thus removing the space dependence (α dependence) and leaving $\Delta g = \text{constant}$. To do this, set the term in parentheses in eq.(2.18) to zero and solve for *w* to get

$$w = \frac{\mu_{HE}^{G} - \mu_{HE}^{A}}{\mu_{LE}^{G} - \mu_{LE}^{A}} \tag{2.19}$$

in which case eq.(2.19) and (2.18) imply

$$\Delta g = g_{HE} - w \cdot g_{LE} = t\mu_{HE}^{A} - wt\mu_{LE}^{A} \tag{2.20}$$

which is a spatially constant featureless image. Again the above derivation works for the monoenergetic case and there will be some residual structure in Δg for a realistic case. Evaluating *w* at average energies, *w* has a numerical value of about 0.3 (Hu and Zhao, 2012).

We can modify the above to include contrast by assuming a glandular fraction of α, an iodine fraction β and an adipose fraction $(1-\alpha-\beta)$. Here, β represents a percentage by volume of iodine, which is quite small. With this scheme, we may rewrite eqs.(2.13) and (2.14) as

$$g_{HE} = t\alpha\mu_{HE}^{G} + t(1-\alpha-\beta)\mu_{HE}^{A} + \beta\mu_{HE}^{I} \tag{2.21}$$

$$g_{LE} = t\alpha\mu_{LE}^{G} + t(1-\alpha-\beta)\mu_{LE}^{A} + \beta\mu_{LE}^{I} \tag{2.22}$$

We can, using eqs.(2.21) and (2.22) go through the same algebraic steps used to get eqs.(2.15)-

(2.20) with w defined as in eq.(2.19). The net result, analogous to eq.(2.20), is

$$\begin{aligned} \Delta g &= g_{HE} - w \cdot g_{LE} \\ &= t\mu_{HE}^{A} - wt\mu_{LE}^{A} \\ &\quad -\beta t(\mu_{HE}^{A} - w\mu_{LE}^{A}) \\ &\quad +\beta t(\mu_{HE}^{I} - w\mu_{LE}^{I}) \end{aligned} \tag{2.23}$$

In eq.(2.23), the first term, $t\mu_{HE}^{A} - wt\mu_{LE}^{A}$, is spatially constant but the 2nd and 3rd terms, $\beta t(\mu_{HE}^{A} - w\mu_{LE}^{A})$ and $\beta t(\mu_{HE}^{I} - w\mu_{LE}^{I})$, vary spatially since β varies with each ray. With $w \approx 0.3$ and the attenuation values read from Figure 2.8 at average LE and HE energies, it is easily seen that the 3rd term in eq.(2.23), corresponding to the iodine image, is about two orders of magnitude greater than the 2nd term, so the 2nd term can be ignored. Thus the 1st and 3rd terms correspond to the constant and iodine components of the weighted subtraction projection image.

One could compute Δg at 0° and thus view a contrast-enhanced dual-energy digital mammogram (CE-DE-DM). Alternatively, dual energy projections could be acquired at many angles and then reconstructed to obtain a 3D view of the contrast. This is CE-DE-DBT. In later chapters we will study aspects of CE-DE-DBT in detail.

Figure 2.9 compares a DBT and CE-DE-DBT image. In the DBT image, the tumor is masked by the fibroglandular and adipose anatomical clutter, but in the CE-DE-DBT image, the anatomical background is suppressed and the iodinated mass easily seen.

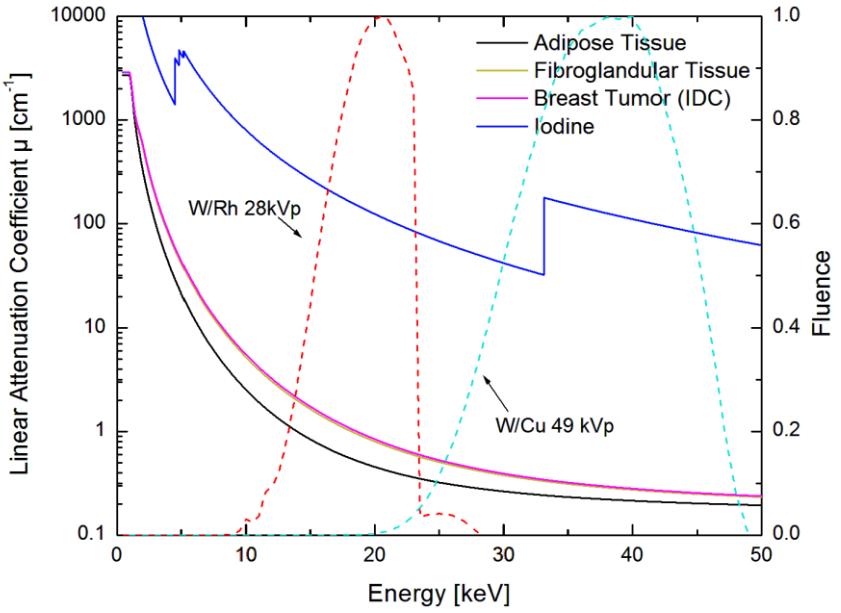

Figure 2.8 Spectra and attenuation coefficients for contrast enhanced imaging. The normalized HE and LE spectra are shown. The attenuation coefficient variation of adipose and fibroglandular breast tissue with energy are shown, as is the relatively high attenuation of iodine. The fibroglandular and tumor attenuation curves overlap. Note the K-edge in the iodine attenuation at 33 keV, within the HE spectrum.

Figure 2.9 Regular DBT vs. CE-DE-DBT. The left image is a conventional DBT slice (MLO view). The tumor is masked by the anatomical clutter caused mostly by adipose and parenchymal tissue variations. The right image is the same slice reconstructed from contrast enhanced

weighted subtraction projection data. The CE-DE-DBT displays the mass prominently. (Images courtesy of Hologic Inc. (Ren et al., 2013))

To do this dual energy subtraction imaging, we need to know several things: what is the dose split to use for high/low energy? That is, if the dose is constrained to an upper limit, what portion should go to LE and what to HE in order to get the best image? The dose split is known approximately but there is still research on this topic, as can be found in (Hu and Zhao, 2012). Secondly, what is the best weight to be used for subtraction? The weight can be derived for the monoenergetic (one HE or LE energy) case and approximated for the more realistic polyenergetic case (Samei and Saunders, 2011).

A crucial problem is that during the wait between high energy and low energy scans, if the patient moves, then the background subtraction is ruined - that is, we get a subtraction of a pattern with a shifted version of itself, which will lead to edge-enhanced artifacts. It destroys the quality of the subtracted image. So it is very important to keep minimal patient motion. The breast is compressed to limit the motion to some extent, but it can only be compressed so much due to discomfort. The time between the high energy scan and low energy scan is already 1-2 minutes. It's long enough for the patient to move and cause problems. One of the things I will address in Chapter 4 is a new acquisition mode that is designed to address this motion artifact. I'll do a contrast-enhanced dual-energy DBT imaging using a phantom to show that our acquisition mode is useful in preventing motion artifacts.

There's a second mode of contrast-enhanced imaging, so called temporal subtraction (Hill et al., 2013). Here, the idea is simpler. One simply takes a high energy scan without iodine, then injects the iodine and takes another high energy image. We then can directly subtract the two images without worrying about dual-energy effects such as the weight between high and low energy. However, the likelihood of patient motion between the two scans in temporal subtraction during the injection is higher than in dual energy subtraction. This is due to the long interscan time because of the injection procedure that takes place between scans. This leaves temporal subtraction susceptible more to subtraction artifacts. For this reason the dual-energy mode has some advantages. In my work, we're only going to consider dual-energy contrast-enhanced DBT imaging. I have given a very qualitative introduction to the field and more details, images and citations will be given in Chapters 3 and 4.

2.8 Scatter psf and SPR

Earlier we mentioned the mechanism of scatter (due to Compton effects) in the interaction of the X-ray beam and the breast tissue. The scatter term is modeled in our imaging equation (2.4). We also gave the basic idea behind CE-DE-DBT (Contrast-Enhanced Dual-Energy DBT) in section 2.7. It turns out that in my work in CE-DE-DBT, scatter will play a major role, so we describe a few more concepts.

Given an X-ray beam, the amount of generated scatter depends on the thickness of material seen by the beam, the energy of the beam and the composition of the material along the beam path. For DBT, the compressed breast is a rectanguloid except at the rounded nipple and lateral edge regions, and the composition is to a great extent partly glandular and partly adipose tissue. Thus for shallow cone parallel-beam X-rays generated from the tube in DBT at angle θ, the breast appears as a constant thickness material.

We can characterize scatter behavior in several ways. One is by the sPSF, the scatter point spread function. For a pencil beam of a particular energy hitting a sample of a particular thickness and composition, the spatial pattern of scattered photons on the detector will be a blob surrounding the point where the primary (unscattered) photons hit. The "blob" is also known as the sPSF. For a finite sized object, such as a compressed breast, the scatter field intensity can fall off due to edge effects. The wider the sPSF, the more pronounced the roll-off effect. For a breast of finite extent, the sPSF is spatially variant.

The sPSF is wider at higher energies. Figure 2.10 shows the sPSF of a 5cm thick breast of 50% glandular / 50% adipose composition at various energies imaged with a 10 mm air gap. The 26 kVp Mo/Mo (i.e. Molybdenum anode and Molybdenum filter) beam generates a sPSF thinner than the 50 kVp (corresponding to HE in CE-DE-DBT) W/Al (Tungsten anode/ Aluminum filter) beam. This increased width of the HE sPSF will have important implications for DE imaging.

In particular, the wider sPSF means that the scatter rolloff at the edge of the object is much more pronounced at HE than at LE. This means, there will be a significant contribution of scattered photons outside of the periphery of the object for the HE case. It will also lead to cupping artifacts, discussed in Chapter 3.

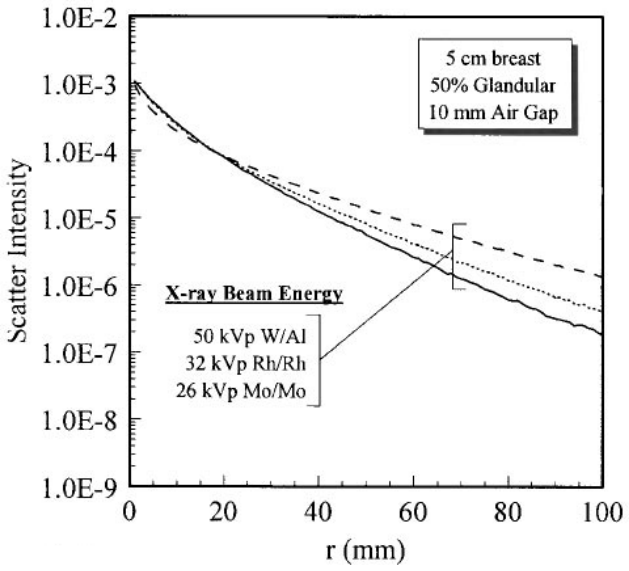

Figure 2.10 Profile of sPSF for a 5cm breast at various energies. (Figure is taken from (Boone and Cooper III, 2000)). The sPSF is evaluated at the center of a phantom using Monte Carlo simulation.

Another entity useful in describing scatter is SPR, the scatter-to-primary ratio. For a given object thickness, composition, and beam spectrum, the SPR at any point on the detector is simply the ratio of the signal from scattered photons to primary photons. Clearly, SPR is object dependent. Nevertheless, the SPR measured at a reference location (the middle of the breast) can be a useful quantity. See (Sechopoulos et al., 2007b) for an extensive discussion of SPR.

In this chapter we have given assorted background material that helps explain the research results presented in Chapters 3 and 4.

Chapter 3 A Scatter Correction Method for Contrast-Enhanced Digital Breast Tomosynthesis

This chapter is an extension of the conference paper (Lu et al., 2014a). A longer, more complete version is currently under review for the journal *Physics in Medicine and Biology* named as " A scatter correction method for contrast-enhanced dual-energy digital breast tomosynthesis ". The work and ideas are mostly mine but I acknowledge the help of co-workers in (Lu et al., 2014a) and "A scatter correction method for contrast-enhanced dual-energy digital breast tomosynthesis".

3.1. Introduction

Conventional X-ray mammography and digital breast tomosynthesis (DBT) rely on attenuation differences to show masses. Masses in conventional mammographic images may be masked by over- and underlying tissue structure, but even in DBT, a mass might not have sufficient contrast to be easily visualized. In this case one can use iodine contrast-enhanced (CE) X-ray imaging, which relies on the fact that, due to angiogenesis, malignant masses contain a dense network of leaky capillaries (Dromain et al., 2012). The contrast material leaks though to capillaries to form, ideally, an easily seen bolus of high-attenuation iodine at the location of the mass (Dromain et al., 2012) (Jong et al., 2003).

Two basic methods for CE breast imaging, temporal subtraction(Skarpathiotakis et al., 2002) and dual-energy (DE) imaging (Puong et al., 2007), have been proposed. We focus on DE here because of its practical acquisition advantages(Puong et al., 2007), and reduction in patient motion artifacts (Carton et al., 2006) . In CE-DE imaging both a diagnostic-range low energy (LE) acquisition, and a high energy (HE) acquisition with mean energy above K-edge of iodine, are acquired. In CE-DE digital mammography (CE-DE-DM), the HE and LE images, acquired at one angle, undergo a weighted subtraction to reveal a 2-D image of the iodine concentration with the anatomical tissue background image suppressed (Lewin et al., 2003). In CE-DE-DBT, the

subtracted images at each angular acquisition serve as an input to a reconstruction algorithm. The resulting 3-D reconstruction shows the iodine concentration with the anatomical background suppressed(Puong et al., 2007). In addition to finding suspicious lesions CE breast imaging has also been applied to obtain contrast kinetics(Dromain et al., 2006), where the kinetic behavior might be used to discriminate benign from cancerous lesions in a manner analogous to that used in Gd-contrast breast MRI.

Conventional breast tomosynthesis has the advantage of displaying a 3-D anatomical image instead of the 2-D projection image of conventional mammography. CE-DE-DBT enjoys this same advantage relative to CE-DE-DM for 3D contrast display. Neither CE-DE-DM nor CE-DE-DBT yet enjoys routine clinical use. Improvements in image quality may further more widespread use. One improvement is effective patient-specific scatter correction. In this paper, we address scatter correction for CE-DE-DBT.

Scatter diminishes iodine visualization and quantification accuracy (Carton et al., 2006, Boone and Cooper III, 2000, Puong et al., 2008, Salvagnini et al., 2012). In particular, for HE acquisitions, scatter can create severe cupping artifacts in the projection data -far more severe than those from diagnostic LE acquisition- that propagate into the reconstruction. Also, without some means of scatter correction (SC), the HE and LE acquisition parameters and weighting factor cannot be optimally selected to suppress anatomical tissue background variations(Carton et al., 2006, Boone and Cooper III, 2000, Puong et al., 2008, Salvagnini et al., 2012).

In order to be useful, the SC computation must be accomplished in a clinically acceptable time. Furthermore, a successful SC technique should not add significant patient acquisitions nor result in the patient spending significant extra time under compression. In addition, a successful SC method must be transportable from one site to another, without a labor-intensive recalibration of site-specific parameters.

A variety of SC methods have been studied for DM and DBT as well as for the related modalities of cone-beam X-ray CT (CBCT) and dedicated breast CT. (Sisniega et al., 2015) used GPU-accelerated Monte Carlo (MC) applied to brain CBCT to rapidly calculate patient-specific scatter estimates. This technique required an initial fully-3D (not limited angle) reconstruction, using uncorrected data followed by anatomical segmentation. MC was then applied to the

segmented image to directly obtain scatter estimates. Processing times of 14 min were reported. Investigators have computed scatter point-spread functions (spsf) (Yang et al., 2014, Lau, 2012, Diaz et al., 2014) to study the nature of scatter in DM and DBT. (Zhao et al., 2015) developed a spsf-based patient specific SC method for CBCT. An initial segmented 3D reconstruction was used to estimate primary data, which could be combined with uncorrected (primary + scatter) data to fit spsf parameters and perform SC. Processing times were about 15 min. (Yang et al., 2014) used a hardware approach for SC in dedicated breast CT. In addition to ordinary projection acquisitions, extra projections utilizing a pinhole array positioned at the X-ray output were used to obtain projections of primary-only data. The two types of projections were combined to obtain a scatter estimate. In another approach used for DBT patient-specific SC (Sechopoulos et al., 2007b, Feng and Sechopoulos, 2011, Feng et al., 2014) precomputed, using offline MC, an extensive "library" of scatter maps for a variety of breast shapes and thicknesses. This method required an extensive library of (about 10^5 entries) and complex image analysis techniques to match a given breast to a library entry to obtain the scatter estimate. An empirical method (Siewerdsen et al., 2006) was applied to CBCT. Here the tails of projection data due to scatter outside the object shadow were interpolated into the shadow region to obtain a scatter estimate.

We propose a SC method intended for clinical use in CE-DE-DBT. Initial results were presented in Lu et al.(Lu et al., 2014a). Our method relies on interpolating the projection image in a region outside the breast shadow into a scatter estimate within the breast shadow. This has some similarities to the method in (Siewerdsen et al., 2006) mentioned above. However, that method was proposed for cone beam CT and the scatter correction problem there differs from ours in many respects.

Though empirical, our SC correction is patient specific and very fast (about 5 minutes but easily accelerated to well under a minute). No extra scans or extra patient dose is required. No extra hardware or hardware development is needed. The SC package can be easily migrated to different sites, requiring only easily obtainable site specific calibration data. Interestingly, interpolation-based empirical SC is routinely used in clinical SPECT. In SPECT "window-based" SC, scatter in a photopeak energy window is estimated by interpolation of scatter data from adjacent satellite energy windows (Ogawa et al., 1991).

In section 3.2, we present the SC algorithm and a methodology for validating the scatter estimate. In section 3.3, we present the validation results for a variety of phantoms. We also apply SC to phantom acquisitions to demonstrate decupping and the removal of iodine signal dependence on breast thickness. Anecdotal clinical results are shown in section 3.4. In section 3.5 we present a discussion of results and limitations of our method.

3.2 Methods

3.2.1. Acquisition details

Testing and validation were accomplished through phantom acquisitions. All acquisitions were done on a Siemens Mammomat Inspiration CEDET DBT prototype unit shown in Figure 3.1. The 300μm a-Se detector has 3584 by 2816 square pixels of dimension 85μm on a side. The X-ray focal-spot to detector distance is 655.5 mm and the center of rotation is located 47 mm above the detector, with air gap between the carbon fiber detector cover and detector of 17 mm. The compression paddle is composed of 3 mm PMMA and its area covers the detector. Images were acquired at 25 angular orientations over an angular span of $\theta_{initial} = -23°$ to $\theta_{final} = +23°$ in steps of $\Delta\theta = 2°$. The X-ray tube motion was continuous.

Figure 3.1 Siemens Mammomat Inspiration CEDET scanner used for acquisitions. A CIRS BR3D (model 020) phantom is shown mounted in the scanner. The inset shows a customized slab of the phantom with iodine inserts.

We used the following parameters for dual-energy phantom acquisition that we previously developed (Hu et al., 2013). The tube parameters were 28 kVp for LE and 49 kVp for HE. The particular mAs used depended on the given experiment, and we report this in section 3.3. We used a W/Rh anode/filter combination for LE and W/Cu for HE. The filter thickness for Rh was 50 μm and for Cu was 237 μm.

For CE-DE-DBT studies, we first used a CIRS BR3D (model 020) phantom. (For convenience, we shall refer to this as a "CIRS-020" phantom.)The CIRS-020 phantom comprises six 1 cm thick semicircular slabs of radius 9 cm containing a mixture of simulated adipose (50%) and glandular (50%) tissue to mimic the human breast. The adipose and glandular components are spatially mixed into a "swirl" pattern as seen in Figure 3.2(a). These slabs can be stacked to obtain a range 2-6 cm of breast thicknesses. One customized slab contained iodine inserts arranged in a 4x4 array as shown in Figure 3.2. As shown in Figure 3.2, the iodine cylinders in each column are 2, 3, 5 and 8 mm in diameter with the depth equal to the diameter. The signals are spaced by 1 cm. The tops of all 16 cylinders lie in one plane. The iodine concentration in each row is (top to bottom) 1, 2, 3 and 5 mg/ml. The thickness of our reconstructed slice was 1 mm and the slice was chosen to intersect all iodine cylinders.

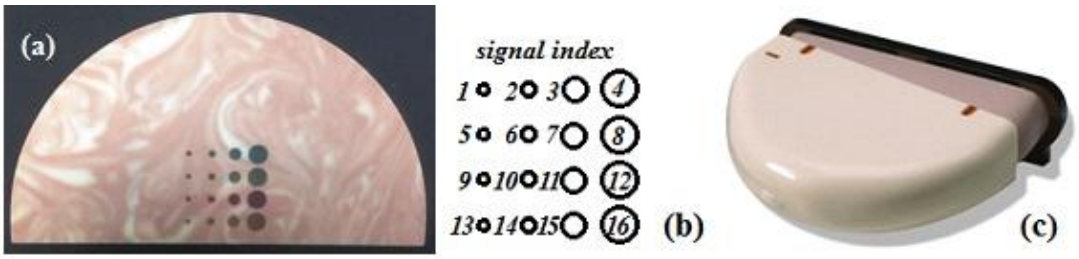

Figure 3.2 CIRS-020 and CIRS-011A phantoms. (a) Photo of 9 cm radius slab with iodine inserts. (b) Drawing to scale of signal region of CIRS-020 shows the iodine insert indices. (c) The CIRS-011A phantom comprises a skin tissue cover (as shown in white) and breast mimicking tissue component (as shown in pink).

Since real breasts under compression have rounded edges and the CIRS-020 has sharp edges, we verified that our scatter correction was unaffected by rounded edges through use of a CIRS tissue-equivalent phantom (model 011A)(CIRS), shown in Figure 3.2 (c), designed to mimic a 4.5 cm thick compressed breast. The lateral dimensions are 12.5 (chest to nipple) and 18.5 cm (width). (For convenience, we shall refer to this as the "CIRS-011A" phantom.) The CIRS-011A has no iodine contrast inserts. It is of uniform 50-50 adipose-glandular composition except for linear fibers, dots (calcifications), cylinders (masses) and calibrated resolution and attenuation reference inserts, but these have no effect on our validations.

3.2.2. Imaging and reconstruction model for CE-DE-DBT

Figure 3.3 illustrates a stylized (not to scale) DBT system with a stationary detector and an X-ray tube that rotates (indicated by angle θ) in the x-z plane with the y coordinate perpendicular to the paper. The system is shown at two angular orientations. Profile plots of detector readout along x at a given y are shown for each of the two angular orientations.

Consider the high-energy case with $I_{HE}^{\theta}(x,y)$ the detector readout at (x,y). The subscript HE indicates high energy acquisitions and superscript θ indicates the acquisition angle. We can write the HE acquisition as

$$I_{HE}^{\theta}(x,y) = b_{HE}^{\theta}(x,y)\exp(-\int_{l_{xy}^{\theta}} \mu_{HE}(x,y,z)dl) + s_{HE}^{\theta}(x,y) + n_{HE}^{\theta}(x,y), \quad (3.1)$$

where b_{HE}^{θ} is the high-energy blank scan, s_{HE}^{θ} is high-energy scatter and n_{HE}^{θ} indicates all radiation and detector noise sources from a high energy acquisition, here written as additive term. The beam attenuation is modeled by the exponentiated line integral, where $\mu_{HE}(x,y,z)$ is the high energy linear attenuation coefficient of the breast, the compression paddle and breast support; *dl* indicates line integral summation with $l_{x,y}^{\theta}$ denoting that the line integral is along a ray from the X-ray focal spot (here considered a point source) to the pixel centered at (x,y). Note that Eq.(3.1) holds for a monoenergetic case, so it can be considered an approximation with HE representing the mean of the HE energy range. Equation (3.1) is written in a continuous (x,y,z) space, but it is understood that (x,y) indicates the center of an integrating detector element centered at (x,y) while for the attenuation object $\mu(x,y,z)$, the coordinates are continuous. Everything follows analogously for the low energy case, where the imaging equation becomes

$$I_{LE}^{\theta}(x,y) = b_{LE}^{\theta}(x,y)\exp(-\int_{l_{xy}^{\theta}} \mu_{LE}(x,y,z)dl) + s_{LE}^{\theta}(x,y) + n_{LE}^{\theta}(x,y) \tag{3.2}$$

with the subscript LE indicating low energy.

Consider the problem of reconstructing the HE data. Typically, a log step of the form

$$g_{HE}^{\theta}(x,y) = -\log\left(\frac{I_{HE}^{\theta}(x,y)}{b_{HE}^{\theta}(x,y)}\right) \tag{3.3}$$

is first applied. If scatter and noise are absent in Eq.(3.1), then $g_{HE}^{\theta}(x,y)$ is simply the line integral data of the object $\mu_{HE}(x,y,z)$. For reasons explained below, we will use the set of $g_{HE}^{\theta}(x,y)$ images (one per θ) as input to a reconstruction algorithm to obtain a 3-D HE reconstruction. However, if scatter is present in eq.(3.1), the detector data $I_{HE}^{\theta}(x,y)$ will contain a "cupping" profile easily seen in Figure 3.3. The cupping is propagated into $g_{HE}^{\theta}(x,y)$ via the log step in eq.(3.3) and is further propagated into the reconstruction such that a reconstructed slice shows a cupping profile. Thus if one could estimate $s_{HE}^{\theta}(x,y)$ and subtract it from $I_{HE}^{\theta}(x,y)$, the subsequent reconstruction would be "decupped" yielding better iodine image quality.

For LE data, the cupping in $I_{LE}^{\theta}(x,y)$ is negligible relative to that in $I_{HE}^{\theta}(x,y)$. Nevertheless, the scatter $s_{LE}^{\theta}(x,y)$, when propagated through the log step and into an LE reconstruction, causes

quantitative errors. Thus estimation of $s_{LE}^{\theta}(x,y)$ and subtraction from eq. (3.2) yields a quantitatively improved reconstruction.

Above we discussed separate reconstructions of the HE and LE data, which differs from the conventional strategy of weighted subtraction-then-reconstruction strategy. Conventionally, in CE-DE-DBT, a weighted subtraction of the form

$$\Delta g^{\theta}(x,y) = g_{HE}^{\theta}(x,y) - W \cdot g_{LE}^{\theta}(x,y) \tag{3.4}$$

is performed, where $\Delta g^{\theta}(x,y)$ is weighted subtracted projection data at θ and W is a weight factor. This yields $\Delta g^{\theta}(x,y)$, a projection image at θ, where ideally, the contrast between adipose and glandular tissue is zero, leaving only an iodine projection image (Samei and Saunders, 2011). The $\Delta g^{\theta}(x,y)$ can then be reconstructed to ideally show iodine contrast only. In practice, polyenergetic effects and patient motion lead to an incomplete subtraction of the breast structure which leads to residual anatomical structure in the CE-DE-DBT reconstruction. The noise will propagate from the projection data to yield correlated noise in the CE-DE-DBT reconstruction. Scatter results in the "cupping" artifact, mainly due to cupping in HE data, that can greatly reduce visibility of the iodine signal in the CE-DE-DBT reconstruction.

We modify the conventional strategy in eq.(3.4) by separately reconstructing g_{HE}^{θ} and g_{LE}^{θ} and then performing the weighted subtraction in the reconstruction domain (Chen et al., 2013). The reconstruction algorithm used in this study was OS-SART (Wang and Jiang, 2004). We used 4 iterations. (We avoided the use of filtered backprojection algorithms since they yield artifacts (Lu et al., 2014b, Chen et al., 2013) in the reconstruction of the iodine signals.) Since OS-SART is linear, it can be shown that the choice of weight is unaffected. The advantage (Chen et al., 2013) of our reconstruct-then-subtract strategy is that it minimizes artifacts due to inexact angular positioning of the X-ray tube. Scatter estimates were subtracted from the projection data (before the log step) before the reconstruction.

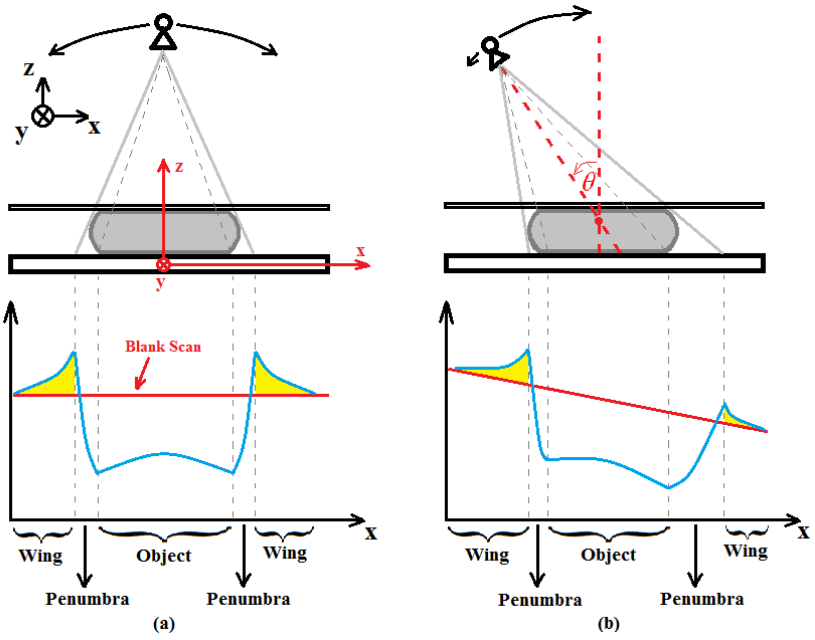

Figure 3.3 Stylized DBT system (not to scale). The x-y-z coordinates are defined in (a). The x-y plane is coincident with the detector plane. A breast sits atop the detector and is irradiated by an X-ray tube that rotates by θ in the x-z plane. The center-of -rotation sits above the x-y plane as shown in (b). Configurations at zero degrees (a) and at an oblique angle (b) are shown. Profiles in the x-direction at a fixed y indicate the detector readout in three regions: object, penumbra and wing, defined in the text.

3.2.3. Scatter correction algorithm

3.2.3.1 Overview

With the aid of Figure 3.3, we illustrate the segmentation of each HE detector readout into three regions. Procedures for the segmentation are given in section 3.2.3.2. The segmentation is illustrated along x for a given value of y. The "wing" region is delineated by the X-ray tangent to the breast. In this region, the readout comprises scatter and blank scan. It is easy to subtract the known blank scan b_{HE}^{θ} in the wing region to obtain s_{HE}^{θ} in the wing region, shown as the yellow area in Figure 3.3. The "object" region comprises that part of the detector readout that has received rays passing through the full thickness of the compressed breast. The "penumbra" region is the remaining segment in between the wing and object regions.

We first give a simple and intuitive overview of the HE SC algorithm with the aid of the scatter profiles in Figure 3.4. In section 3.2.3.2, we revisit the SC algorithm in mathematical detail. Note that the LE SC algorithm is a simplified version of the HE SC and will be described in section 3.2.3.2.

In Figure 3.4(a), the true scatter profile is shown in the 3 regions. As can be seen in Figure 3.3, the scatter in the wing region can be obtained by subtraction of the (known) blank scan. Our goal is to obtain an estimate of the true scatter in the object region using knowledge of the wing data along with pre-stored information.

We presume that the profile in the object region is a flattened bell-shaped curve. As a first step, we use polynomial-based interpolation of the wing data to obtain an initial estimate of the scatter in the object and penumbral regions. The interpolated data is shown in Figure 3.4(b). As we will validate, with an appropriate choice of the interpolating polynomial, the shape of the interpolated curve in the object region is nearly correct and differs from the true scatter data by an additive constant k. This feature is illustrated in Figure 3.4(c). We will be able to obtain k using a relatively simple procedure involving pre-stored measurements as detailed in section 3.2.3.2.

We add k to the interpolated scatter estimate in the object and penumbral regions as seen in Figure 3.4(d). In Figure 3.4(d) we see that resulting scatter estimate is fairly accurate in the object region but is a poor approximation of the true scatter in the penumbral region. The poor approximation of scatter in the penumbral region causes artifacts in the dual-energy reconstruction as will be shown in later sections, but these artifacts are insignificant compared to artifacts due to not doing any scatter correction in the first place. The scatter estimate for this simple 1-D case is shown in Figure 3.4(d).

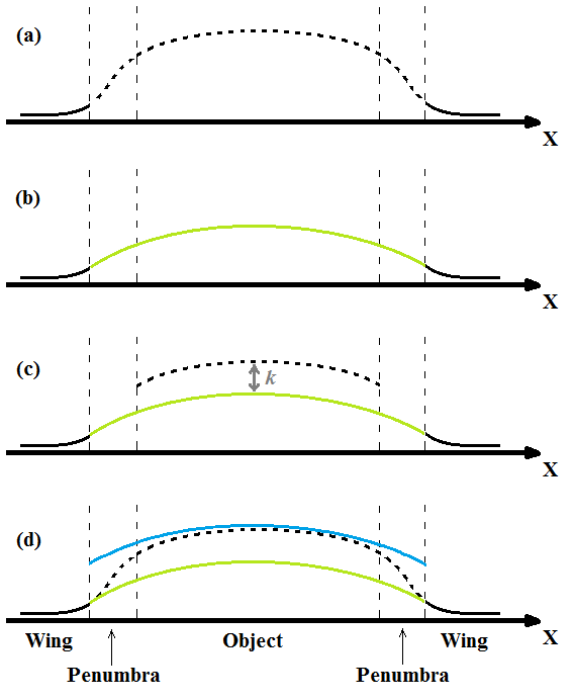

Figure 3.4 Illustration of 1-D version of scatter estimation. (See text for detailed explanation). (a) True scatter profile. (b) Green curve is polynomial interpolation from wing data. (c) In object region, interpolated profile differs from true scatter by additive offset k. (d) Blue curve, obtained by adding k to the interpolated green curve, is the scatter estimate. It is accurate in the object region and an inaccurate in the penumbral region.

In Siewerdsen et al.(Siewerdsen et al., 2006), the notion of scatter estimation by interpolated wing data is used in the context of cone-beam CT. The interpolation in itself is supplemented by an additional empirical step involving truncation of the wing region before interpolation.

3.2.3.2 Details

We describe the SC algorithm in detail with the aid of the table of symbols in Table 3.1 and the pseudocode in Table 3.2. We first consider HE data only. The SC operations are performed at each angle, as indicated by the superscript θ. For a given patient, the HE projection data $I_{HE}^{\theta}(x,y)$ is acquired along with a pre-measured and dose-scaled blank scan $b_{HE}^{\theta}(x,y)$.

We first list the steps used in segmenting and extracting the wing data $w_{HE}^{\theta}(x,y)$ from $I_{HE}^{\theta}(x,y)$ and $b_{HE}^{\theta}(x,y)$. We convolve the blank scan $b_{HE}^{\theta}(x,y)$ with a Gaussian kernel (limited to 15x15 pixel extent and with $\sigma = 7$ pixels) to obtain a smoothed blank scan $\tilde{b}_{HE}^{\theta}(x,y)$. Wing data $w_{HE}^{\theta}(x,y)$ is obtained by the following operation:

$$w_{HE}^{\theta}(x,y) = [(I_{HE}^{\theta}(x,y) - \tilde{b}_{HE}^{\theta}(x,y)) \otimes g(x,y)]_{+} \tag{3.5}$$

The subtraction in eq.(3.5) yields wing data which is further smoothed by convolution ($\otimes$) with a Gaussian kernel $g(x,y)$ restricted to a 5x5 pixel region and with $\sigma = 3$. Any negative values are set to zero as indicated by the clipping operation $[\ ]_{+}$. The subtraction and clipping operations perform the essential segmentation step for wing data since the penumbral and object region data lie below the blank scan as depicted in Figure 3.3. The cyan pixels in Figure 3.5 show an x-profile of $w_{HE}^{\theta}(x,y)$. The segmentation is not perfect; as seen in Figure 3.5(a), there are a few outlier pixels on the inner edges of the HE wing data. These are due to a region in the penumbra adjacent to the wing, where the penumbral values exceed the blank scan. Figure 3.3 illustrates this penumbral effect.

Table 3.1 The subscripts HE and LE refer to high energy and low energy. The superscript θ refers to the angle of acquisition. The coordinate (x,y) is position in the detector plane.

Symbols	Descriptions
$I_{HE}^{\theta}(x,y)$, $I_{LE}^{\theta}(x,y)$	Raw projection data
$g_{HE}^{\theta}(x,y)$, $g_{LE}^{\theta}(x,y)$	Projection data after log step
$b_{HE}^{\theta}(x,y)$, $b_{LE}^{\theta}(x,y)$	Blank scan data
$\tilde{b}_{HE}^{\theta}(x,y)$, $\tilde{b}_{LE}^{\theta}(x,y)$	Smoothed versions of $b_{HE}^{\theta}(x,y)$, $b_{LE}^{\theta}(x,y)$
$w_{HE}^{\theta}(x,y)$	Smoothed wing data of HE acquisition
$s_{HE}^{\theta}(x,y)$, $s_{LE}^{\theta}(x,y)$	True scatter
$\tilde{\tilde{s}}_{HE}^{\theta}(x,y)$	Interpolation of wing data along x for a fixed y (repeated for each y)
$\tilde{s}_{HE}^{\theta}(x,y)$, $\tilde{s}_{LE}^{\theta}(x,y)$	Smoothed (in y direction) version of $\tilde{\tilde{s}}_{HE}^{\theta}(x,y)$. Note $\tilde{\tilde{s}}_{LE}^{\theta}(x,y)$ is set to zero
k_{HE}^{θ}, k_{LE}^{θ}	Constant used to correct $\tilde{s}_{HE}^{\theta}(x,y)$, $\tilde{s}_{LE}^{\theta}(x,y)$
$\hat{s}_{HE}^{\theta}(x,y)$, $\hat{s}_{LE}^{\theta}(x,y)$	Final estimated scatter
$SPR_{HE}^{\theta}(x,y)$, $SPR_{LE}^{\theta}(x,y)$	True scatter-to-primary ratios
$\hat{SPR}_{HE}^{\theta}(x,y)$, $\hat{SPR}_{LE}^{\theta}(x,y)$	Estimates of scatter-to-primary ratio derived using our scatter correction algorithm
$\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta},\bar{y}_c^{\theta})$, $\overline{SPR}_{LE}^{\theta}(\bar{x}_c^{\theta},\bar{y}_c^{\theta})$	Reference library SPR point value obtained from a phantom at location $(\bar{x}_c^{\theta},\bar{y}_c^{\theta})$
$(x_c^{\theta},y_c^{\theta})$	Reference point near center of breast, location where $\tilde{s}_{HE}^{\theta}(x,y)$ peaks
$(\bar{x}_c^{\theta},\bar{y}_c^{\theta})$	Reference point near center of phantom, location where $s_{HE}^{\theta}(x,y)$ peaks
$p_{HE}^{\theta}(x,y)$, $p_{LE}^{\theta}(x,y)$	Primary photon readings

We obtain the interpolated HE scatter profile of Figure 3.4(b) by performing, for each value of y, a 1-D polynomial interpolation between the two wing regions. Our validation studies in section 3.3.1 show that a polynomial of order 4 yields the best results. To implement the 4th order 1-D polynomial interpolation, we used the Matlab 2013a curve fitting function *fit()*. Note that the effect of the outlier pixels on the interpolation can be controlled by selecting a robust interpolation option (LAR) in Matlab. (We could have modified the wing segmentation to eliminate the outlier pixels but we found that robust interpolation was easier and led to equally good results.) The green curve in Figure 3.5(a) shows the resulting interpolated scatter estimate. The interpolation along x for a fixed y is repeated for each y to obtain a surface $\tilde{\tilde{s}}^{\theta}_{HE}(x,y)$ smooth in x but choppy in y as seen in Figure 3.5(b). The "choppiness" stems from the fact that each x interpolation is from a different noisy set (different y) of wing data. Since we know that the scatter field should be smooth in y, we apply a further smoothing in the y direction using a 1-D Gaussian kernel ($\sigma = 60$, 200 pixel extent) to yield a smooth surface $\tilde{s}^{\theta}_{HE}(x,y)$ as illustrated in Figure 3.5(c).

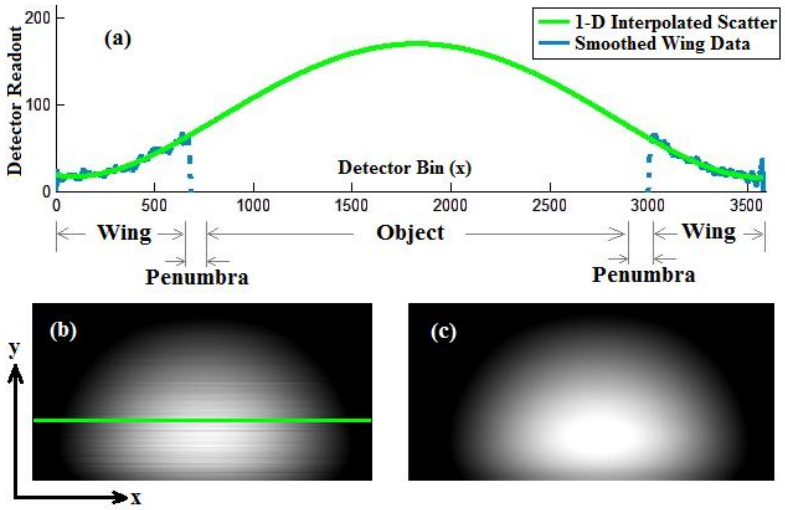

Figure 3.5 Obtaining initial scatter estimate from wing data. In (a), the segmented and smoothed wing data (cyan) and its 4th-order polynomial interpolation (green) are shown for one value of y and for $\theta = 0°$. In (b), the interpolated curves $\tilde{\tilde{s}}^{\theta}_{HE}(x,y)$ for all y are shown as a greyscale image. The green line corresponds to the 1-D profile of scatter in (a). Further smoothing in the y-direction results in the image $\tilde{s}^{\theta}_{HE}(x,y)$ in (c).

We assume (and later validate) that $\tilde{s}^{\theta}_{HE}(x,y)$ will have an approximately correct shape within the object region, but will be quantitatively inaccurate. It underestimates the true scatter $s^{\theta}_{HE}(x,y)$ in the object region by an additive *constant* k^{θ}_{HE} as depicted in Figure 3.4(c), so that

$$k^{\theta}_{HE} = s^{\theta}_{HE}(x,y) - \tilde{s}^{\theta}_{HE}(x,y). \tag{3.6}$$

(Note that k depends on θ and energy.) This underestimate is due to the fact that we cannot accurately interpolate through the rapidly rising scatter in the penumbral region. Figure 3.4(d) depicts this underestimation.

A key step in the SC algorithm is determination of k^{θ}_{HE}. If we assume that eq.(3.6) holds for all the points in the object region, then it must hold at a reference location $(x^{\theta}_c, y^{\theta}_c)$ in the object region:

$$k^{\theta}_{HE} = s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c) - \tilde{s}^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c), \tag{3.7}$$

where $(x^{\theta}_c, y^{\theta}_c) \equiv \arg\max_{x,y}\left[\tilde{s}^{\theta}_{HE}(x,y)\right]$. The true scatter $s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)$ is unknown, but we can approximate it with aid of the scatter-to-primary ratio evaluated at $(x^{\theta}_c, y^{\theta}_c)$:

$$SPR^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c) = \frac{s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)}{p^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)} \equiv \frac{s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)}{I^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c) - s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)}, \tag{3.8}$$

where $p^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)$ is the intensity of the primary (unscattered) rays. Solving eq.(3.8) for $s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)$ yields

$$s^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c) = \left[\frac{SPR^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)}{1 + SPR^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)}\right] I^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c). \tag{3.9}$$

where $I^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)$ is just the patient measurement data at point $(x^{\theta}_c, y^{\theta}_c)$.

We do not know $SPR^{\theta}_{HE}(x^{\theta}_c, y^{\theta}_c)$ for a given patient but we can approximate it using a library of precomputed point measurements obtained from breast-equivalent phantoms as detailed in section 3.2.4. In general, the scatter-to-primary ratio at a central point depends(Sechopoulos et al., 2007b) on thickness t, energy, angle θ and to a lesser extent the glandular portion. The scatter-to-primary ratio measured at $(x^{\theta}_c, y^{\theta}_c)$ for the patient breast will be close in value to the scatter-to-primary ratio measured at a suitably defined central point of a phantom $(\bar{x}^{\theta}_c, \bar{y}^{\theta}_c)$. For a breast

and its phantom equivalent, i.e. the phantom that matches the breast in t and α, we can thus write

$$SPR_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta}) \approx \overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta}) \tag{3.10}$$

where the bar superscript indicates a phantom quantity. Thus our problem is one of determining from phantoms a library of scatter-to-primary *point* measurements, not a library of scatter-to-primary *images*. In section 3.2.4, we detail a procedure for determining $(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ and $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ as well as the same quantities for the LE case.

From eqs.(3.9) and (3.10), we may write

$$\hat{s}_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta}) = \left[\frac{\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})}{1 + \overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})} \right] I_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta}) \text{ ,} \tag{3.11}$$

where $\hat{s}_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta})$ is an estimate of $s_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta})$ due to the approximation in Eq.(3.10). We use eqs. (3.7) and (3.11) to obtain

$$k_{HE}^{\theta} \approx \hat{s}_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta}) - \tilde{s}_{HE}^{\theta}(x_c^{\theta}, y_c^{\theta}) . \tag{3.12}$$

From eq.(3.6), eq.(3.12) holds at all (x, y) in the object region, so that

$$\hat{s}_{HE}^{\theta}(x, y) \approx \tilde{s}_{HE}^{\theta}(x, y) + k_{HE}^{\theta} . \tag{3.13}$$

where $\hat{s}_{HE}^{\theta}(x, y)$ is our final scatter estimate. This estimate can be subtracted from eq.(3.1) to perform scatter correction.

To obtain the low energy scatter estimate in the object region $\hat{s}_{LE}^{\theta}(x, y)$, we follow the steps for the HE case but set $\tilde{s}_{LE}^{\theta}(x, y) = 0$. The reason for this is that the wing data for the LE case is essentially zero due to the narrowness of the LE spsf. Within the object region we approximate $\hat{s}_{LE}^{\theta}(x, y)$ as a constant k_{LE}^{θ}. Numerous studies(Sechopoulos et al., 2007b, Feng et al., 2014, Boone and Cooper III, 2000) have shown a slow spatial variation of low energy scatter in the object region, but we use a constant because the LE spatial variation is so much smaller than the HE scatter variation that the constant approximation yields good results. The remaining steps follow as in the HE case. In particular, we build a library $\overline{SPR}_{LE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ using LE acquisitions of phantoms of varying thicknesses and use these values in eq.(3.11).

The SC algorithm for HE is summarized in the pseudocode of Table 3.2. The LE SC algorithm, described above, is a simplification of the HE algorithm. For convenience, the term

$G_{15\times15}^{\sigma=7}(x,y)$ refers to a 2-D Gaussian smoothing kernel with standard deviation 7 pixels and extent limited to 15×15 pixels. This construct also appears in the kernels $G_{5\times5}^{\sigma=3}(x,y)$ and the 1-D kernel $G_{1\times200}^{\sigma=60}(y)$.

Figure 3.6 illustrates the x-profiles of all processing stages of the algorithm for HE acquisitions using a CIRS-020 phantom at $\theta=0°$ and $\theta=23°$ in place of a real breast.

Table 3.2 Pseudocode for HE SC Algorithm.

---

for θ = $\theta_{initial}$ to θ_{final} in steps of $\Delta\theta$

$\tilde{b}_{HE}^{\theta}(x,y)=b_{HE}^{\theta}(x,y)\otimes G_{15\times15}^{\sigma=7}(x,y)$

$w_{HE}^{\theta}(x,y)=\left[(I_{HE}^{\theta}(x,y)-\tilde{b}_{HE}^{\theta}(x,y))\otimes G_{5\times5}^{\sigma=3}(x,y)\right]_{+}$ ($[\ ]_{+}$ means clip negative values to 0)

for y_{ind} = 1 to (# of detector bins in y)

1-D 4th order polynomial interpolation through $w_{HE}^{\theta}(x,y_{ind})$ to compute $\tilde{\tilde{s}}_{HE}^{\theta}(x,y_{ind})$

end

for x_{ind} = 1 to (# of detector bins in x)

$\tilde{s}_{HE}^{\theta}(x_{ind},y)=\tilde{\tilde{s}}_{HE}^{\theta}(x_{ind},y)\otimes G_{1\times200}^{\sigma=60}(y)$

e**nd**

$(x_c^{\theta},y_c^{\theta})=\arg\max(\tilde{s}_{HE}^{\theta}(x,y))$

Look up $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta},\bar{y}_c^{\theta})$ from library

$$\hat{s}_{HE}^{\theta}(x_c^{\theta},y_c^{\theta})=I_{HE}^{\theta}(x_c^{\theta},y_c^{\theta})\cdot\frac{\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta},\bar{y}_c^{\theta})}{1+\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta},\bar{y}_c^{\theta})}$$

$k_{HE}^{\theta}=\hat{s}_{HE}^{\theta}(x_c^{\theta},y_c^{\theta})-\tilde{s}_{HE}^{\theta}(x_c^{\theta},y_c^{\theta})$

$\hat{s}_{HE}^{\theta}(x,y)=\tilde{s}_{HE}^{\theta}(x,y)+k_{HE}^{\theta}$ is the final scatter estimate

end

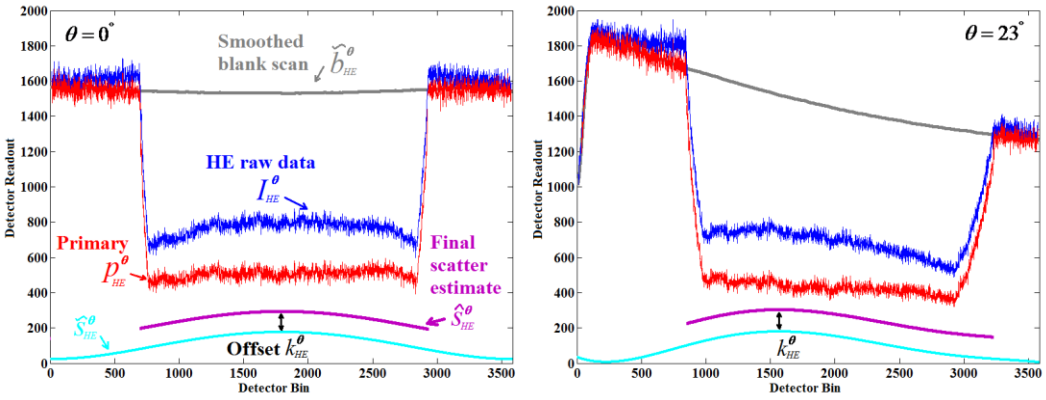

Figure 3.6 Profiles in the x-direction of the various stages of the SC algorithm applied for HE acquisition of a CIRS-020 phantom. Results are shown for two angles. The algorithm takes the raw data I_{HE}^{θ} (in blue) and ends with the scatter corrected data p_{HE}^{θ} (in red).

3.2.4. Determining pre-measured library quantities

From section 3.2.3.2, the SC algorithm requires knowledge of $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ and $\overline{SPR}_{LE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$, the scatter-to-primary ratios measured from reference phantoms. Again, we discuss measurements in terms of HE, but this will apply equally to the LE case. The quantities $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ comprise a library of scalars indexed by θ and phantom thickness t. (In the notation for the library quantity $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ we suppress t.) For a given patient, we use the measured compressed breast thickness to find the corresponding value of t. The thickness t should ideally equal the thickness of compressed patient breast. For our own work, we have available phantoms of thicknesses 2, 3, 4, 5 and 6 cm obtained by stacking 1 cm thick CIRS-020 slabs. We interpolate the $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ values to match the thickness of the patient breast. We note that scatter-to-primary ratio also varies with the glandular/adipose composition ratio, but this variation is slow(Sechopoulos et al., 2007b). The glandular/adipose ratio for the CIRS-020 is fixed at 50%, but given the weak dependence of the scatter-to-primary ratio on this compositional ratio, we do not incur a significant error in $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$.

We will need to measure a scatter-to-primary ratio measurement only in the central region about $(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ of the phantom corresponding to a central region about $(x_c^{\theta}, y_c^{\theta})$ of the

compressed breast. Since the scatter-to-primary ratio at a central location is insensitive to the overall breast shape (Sechopoulos et al., 2007b), the shape of the phantom need not match the shape of the patient breast. The point (x_c^θ, y_c^θ) is the location where $\tilde{s}_{HE}^\theta(x,y)$ attains a maximum, which in turn is the location where $\hat{s}_{HE}^\theta(x,y)$ attains a maximum by virtue of eq.(3.13). Since $\hat{s}_{HE}^\theta(x,y)$ is approximately equal to $s_{HE}^\theta(x,y)$, (x_c^θ, y_c^θ) is the approximate location of maximum scatter in the patient data. We therefore measure $\overline{SPR}_{HE}^\theta(\bar{x}_c^\theta, \bar{y}_c^\theta)$ of the reference phantom at a point $(\bar{x}_c^\theta, \bar{y}_c^\theta)$ where scatter is a maximum. To find this maximum, we must measure the phantom scatter-to-primary ratio at an array of locations. We can obtain these measurements using techniques that entail the use of a pinhole array(Chen et al., 2012, Inscoe et al., 2013, Yang et al., 2014). Details of our adaptation and use of the pinhole array measurement technique (Chen et al., 2012, Inscoe et al., 2013, Yang et al., 2014) are found in section 3.2.5 as part of our description of validation techniques. We define, for a given θ, $(\bar{x}_c^\theta, \bar{y}_c^\theta)$ as that pinhole array shadow location that is within the object region of the breast shadow and also has the maximum scatter-to-primary ratio. This scatter-to-primary ratio value is our $\overline{SPR}_{HE}^\theta(\bar{x}_c^\theta, \bar{y}_c^\theta)$. By performing the procedure at HE and LE, we obtain $\overline{SPR}_{HE}^\theta(\bar{x}_c^\theta, \bar{y}_c^\theta)$ and $\overline{SPR}_{LE}^\theta(\bar{x}_c^\theta, \bar{y}_c^\theta)$ and use these quantities in eq.(3.11).

3.2.5. Pinhole array technique for scatter-to-primary ratio measurement

The pinhole array method can deliver measurements of primary-only photons for a given phantom. Below we describe our use of this technique and also cross-check its ability to deliver primary-only measurements with the aid of a beam-blocker technique (Segui and Zhao, 2006, Niu and Zhu, 2011, Lazos and Williamson, 2012). We describe the technique for the HE case only, but the steps are the same for the LE case.

We place a lead plate with a pinhole atop a 3.9 cm thick uniform Lucite phantom. The plate thickness is 3 mm and hole diameter 1 mm. The lead plate mounted on the scanner is tilted by θ so that it is perpendicular to the X-ray beam. At any θ the plate is positioned so the focal-spot-to-hole distance remains constant. The pinhole shadow will comprise only primary photons as we will verify below. The plate is then removed and the exposure repeated. Now the photons in

the pinhole shadow region comprise total (scatter + primary) photons. From the pinhole shadow, we obtain a measure of the primary photons p by averaging a 5x5 pixel region at the center of the approximately circular shadow. With the plate removed, we can obtain a measure of total counts $I = p + s$ by averaging over the same region. The two quantities I and p can be subtracted to obtain the scatter s. Thus from the two acquisitions, we can obtain the scatter s and the primary p for the phantom at the pinhole shadow. From the two measurements, we can thus compute the true scatter-to-primary ratio $SPR = s / p$ within the pinhole shadow region. The acquisitions can be done at non-clinical high mAs to obtain nearly noiseless quantities. We can obtain SPR at all θ by tilting the plate. The location of the pinhole shadow will vary with θ. The HE acquisitions with and without the plate are acquired at 36 mAs per angle and for the LE case 54 mAs per angle.

To verify that for the given pinhole dimensions the reading in the shadow region comprises only primary photons, we performed the following cross-check. We used a beam-blocker method(Segui and Zhao, 2006, Niu and Zhu, 2011, Lazos and Williamson, 2012) with blocking disks of 2mm - 7 cm for the case of $\theta = 0°$ and for HE and LE exposures equal to that used for the single-pinhole case. The beam blocker method will deliver a scatter value s_{BB} at the pinhole shadow location. The pinhole method delivers $I - p = s$. Our cross -check involved comparing s_{BB} with s. The difference with the two methods was less than 0.5%. In order to reduce discrepancies in s_{BB} vs. s due to phantom inhomogeneity, we used the uniform Lucite phantom. With the uniform phantom, exact positioning of the pinhole versus the blocking disks was not critical.

The single-pinhole experiment above was used to verify that we could obtain s and p in the pinhole shadow. However, to determine $(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ and $\overline{SPR}_{HE}^{\theta}(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$ we need an array of pinholes. We use a 3 mm lead plate oriented at θ as before atop the Lucite phantom, but now the lead plate contains an 15x15 array of pinholes arranged in a square grid with hole centers spaced at 1 cm. Figure 3.7 shows the pinhole placement along with a picture of the lead pinhole plate. The HE and LE acquisition parameters were the same as for the single-pinhole case.

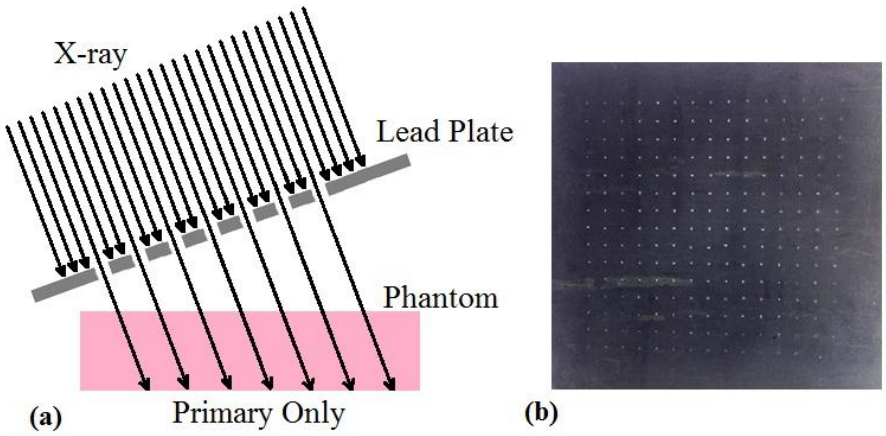

Figure 3.7 Pinhole array acquisition (a) Diagram (not to scale) shows acquisition with plate oriented at θ. The counts in the pinhole shadow will comprise primary-only photons. (b) The lead plate with a 15x15 array of holes.

The photons received in each of the 225 shadow regions constitute primary photons $p_{HE}^{\theta}(x,y)$. With the plate removed, we receive photons $I_{HE}^{\theta}(x,y) = p_{HE}^{\theta}(x,y) + s_{HE}^{\theta}(x,y)$ in the 225 shadow regions; by subtraction, we obtain $s_{HE}^{\theta}(x,y)$ separately. Within each shadow region, we average a 5×5 pixel region at the center of the shadow to obtain scalars $s_{HE}^{\theta}(x,y)$ and $p_{HE}^{\theta}(x,y)$ where (x,y) is now understood to be the pixel location at the center of the shadow. Then we may set $SPR_{HE}^{\theta}(x,y) \equiv \frac{s_{HE}^{\theta}(x,y)}{p_{HE}^{\theta}(x,y)}$ as the scatter-to-primary ratio at each pinhole shadow. We use the same procedure for LE to obtain $SPR_{LE}^{\theta}(x,y)$. Figure 3.8 shows a reading with plate present and absent over a region that includes three pinhole shadows.

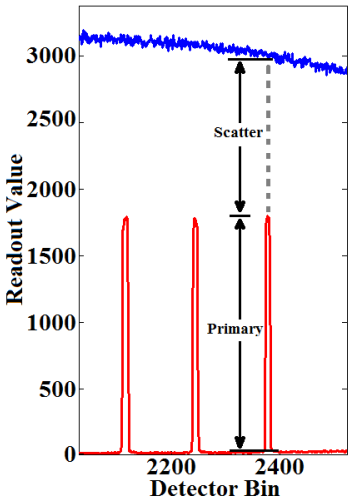

Figure 3.8 Partial profile of image of a Lucite phantom HE acquisition with (a) plate inserted (red), showing the image due to three adjacent pinholes and (b) with plate removed (blue). For a given hole region, "Primary" indicates the contribution due to primary photons and "Scatter" due to scattered photons.

The pinhole-array method requires that contributions through one pinhole do not add counts to other pinhole shadows. To verify that pinholes do not interact, we use the Lucite phantom and acquire acquisitions with plate present and absent as described previously. We measure contributions at θ for all 225 holes and, by using a second plate to cover all holes except one reference hole, measure the contribution due to the reference hole. Figure 3.9 illustrates superposed profiles from the two measurements. Qualitatively, it is seen in Figure 3.9 that the holes do not interact. For a quantitative comparison, we computed the mean m_1 of a 5×5 pixel region in the reference hole shadow with all other holes blocked and a similar mean m_2 of the pixels in the reference hole shadow with all other holes open. We found that $\frac{m_2-m_1}{m_1}<0.01$. Note that m_1 and m_2 are essentially deterministic due to the very high dose of the measurements. i.e. the standard deviations of m_1 and m_2 from repeated measurements are less than m_2-m_1. We repeated this procedure for two more reference holes and obtained similar results.

Our application of the pinhole array technique extends the work in (Chen et al., 2012, Inscoe et al., 2013) in two ways. First, we used a beam blocker method to cross-check the primary readings as describe above, and we also applied this technique in a geometry where the X-ray tube and pinhole plate were positioned at angle θ relative to the detector.

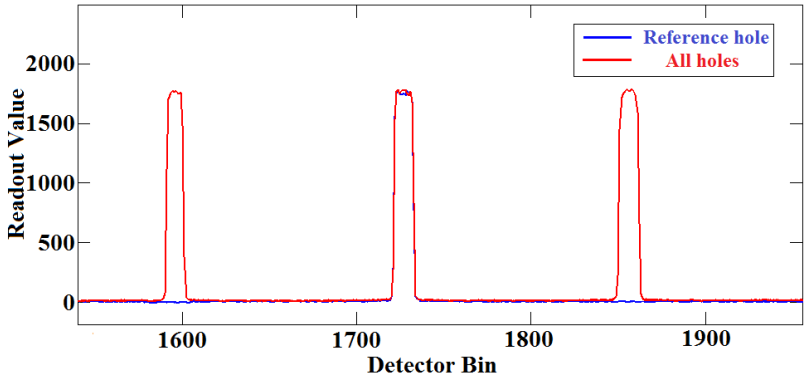

Figure 3.9 Test of pinhole shadow interactions. The red curve is a profile of 3 of the 225 simultaneously irradiated pinholes and the blue curve the profile obtained from irradiating the reference hole only. The left and right peaks are from pinholes that neighbor the reference pinhole. See text for a quantitative analysis.

Exposures with and without a pinhole array can be used to directly compute scatter at dense array of points in a given projection. Thus one might simply mount a pinhole plate on the X-ray output during clinical exposure and compute scatter without going through the many empirical steps we propose. Indeed, for dedicated breast-CT, (Yang et al., 2014) proposed exactly this technique, though the number of angles for which the pinhole plate was used is limited. In our work, we avoid the strategy of directly using a pinhole plate since it entails extra patient exposure and also requires hardware development for each new DBT site. However, we do indeed use the pinhole plate for validation purposes as described in the next section.

3.2.6. Validation of scatter correction algorithm

We describe our validation for the HE case; the LE case follows. In section 3.2.5, we showed how we can use pinhole array techniques to obtain ground-truth values $SPR^{\theta}_{HE}(x,y)$ at the pinhole shadow locations. At these locations, we compare $SPR^{\theta}_{HE}(x,y)$ to $\hat{SPR}^{\theta}_{HE}(x,y)$, where $\hat{SPR}^{\theta}_{HE}(x,y)$ is obtained using the scatter estimate $\hat{s}^{\theta}_{HE}(x,y)$ of eq.(3.13) as follows:

$$\hat{SPR}^{\theta}_{HE}(x,y) \equiv \frac{\hat{s}^{\theta}_{HE}(x,y)}{\hat{p}^{\theta}_{HE}(x,y)} = \frac{\hat{s}^{\theta}_{HE}(x,y)}{I^{\theta}_{HE}(x,y) - \hat{s}^{\theta}_{HE}(x,y)}. \tag{3.14}$$

One might propose comparing our estimate $\hat{s}_{HE}^{\theta}(x,y)$ to a ground truth $s_{HE}^{\theta}(x,y)$ obtained using the pinhole-array technique. However, $\hat{s}_{HE}^{\theta}(x,y)$ is naturally estimated from data acquired at clinical exposure levels while $s_{HE}^{\theta}(x,y)$ must be acquired with non-clinical high exposure levels. The amount of scatter is exposure-dependent, so $\hat{s}_{HE}^{\theta}(x,y)$ and $s_{HE}^{\theta}(x,y)$ are incommensurate. However, the scatter-to-primary ratio $SPR_{HE}^{\theta}(x,y)$ and $\hat{SPR}_{HE}^{\theta}(x,y)$ are exposure independent, and thus are commensurate and can be used for validation. Note that the pinhole acquisitions are thus used for two purposes: to obtain the library data of section 3.2.5 and for validation.

Naturally, $SPR_{HE}^{\theta}(x,y)$ and $\hat{SPR}_{HE}^{\theta}(x,y)$ should be derived from data on the same object. In particular, one would want to ideally compare $SPR_{HE}^{\theta}(x,y)$ and $\hat{SPR}_{HE}^{\theta}(x,y)$ on patient data, but this is wholly impractical since this would entail multiple high-dose patient exposures with and without the pinhole plate mounted on the rotating scanner. Instead we compare $SPR_{HE}^{\theta}(x,y)$ and $\hat{SPR}_{HE}^{\theta}(x,y)$ for the CIRS-020 phantom.

We performed validations, i.e. comparing $\hat{SPR}_{HE}^{\theta}(x,y)$ to $SPR_{HE}^{\theta}(x,y)$ and $\hat{SPR}_{LE}^{\theta}(x,y)$ to $SPR_{LE}^{\theta}(x,y)$, using a CIRS-020 phantom with thicknesses 2,3,4,5 and 6 cm. A potential problem with using this phantom in that it has sharp edges, thus leading to possibly unrealistic penumbral and wing regions when compared to the rounded-edge of a real compressed breast. To address this, we performed another validation using the CIRS-011A phantom, a phantom with rounded edges designed to emulate a compressed breast (see Figure 3.2 (c)). All validation results are reported in section 3.3.

For both phantom validation studies, in obtaining $\hat{SPR}_{HE}^{\theta}(x,y)$ and $\hat{SPR}_{LE}^{\theta}(x,y)$, we used 66 mAs over 25 angles for HE acquisitions and 66-195 mAs for LE acquisitions, with the LE mAs varying with phantom thickness. For HE, we were limited to 66 mAs, a value approximately in a clinical range for a 4 cm thick breast. Using an exposure above 66 mAs on our Siemens Mammomat system resulted in undesirable detector saturation in the wing regions. For both phantoms and both energies and for all angles, $SPR_{HE}^{\theta}(x,y)$ and $SPR_{LE}^{\theta}(x,y)$ were calculated at

3x the dose used for $\hat{SPR}^{\theta}_{HE}(x,y)$ and $\hat{SPR}^{\theta}_{LE}(x,y)$. We repeated acquisitions three times and took the average.

3.2.7. Effects of scatter correction on reconstructed image quality

Our scatter correction method was tested on a CIRS-020 phantom with one slab containing iodine inserts. We acquired HE and LE images and performed an OS-SART reconstruction as described in section 3.2.2. A reconstructed slice was used to test the visibility of iodinated lesions with and without our scatter correction method. The 1 cm thick slab with inserts was always positioned is the second slab from the bottom.

We evaluated for the CIRS-020 phantom the degree of decupping afforded by SC, using a metric(Siewerdsen et al., 2006)

$$t_{cup} = 100\frac{\bar{\mu}_{edge} - \bar{\mu}_{center}}{\bar{\mu}_{edge}} \tag{3.15}$$

where $\bar{\mu}_{edge}$ is a pixel average from five regions along the periphery of the reconstructed image and $\bar{\mu}_{center}$ a pixel average in a center region. Detailed results are given in section 3.3.3.

We report in section 3.4 anecdotal results of a CE-DE-DM patient image acquired under mammographic acquisition with a CC view. The images with and without our scatter correction are shown. The patient had a biopsy-confirmed tumor that was clearly visible in the image with SC but invisible without SC.

3.3. Results

In section 3.3.1 we validate our scatter correction algorithm on the CIRS-020 phantom by comparing $\hat{SPR}^{\theta}_{HE}(x,y)$ to $SPR^{\theta}_{HE}(x,y)$ (and $\hat{SPR}^{\theta}_{LE}(x,y)$ vs. $SPR^{\theta}_{LE}(x,y)$) at selected locations, for a variety of angles and a range of thicknesses. In section 3.3.2 this validation is repeated for the rounded edge CIRS011A phantom. In section 3.3.3 we show HE, LE projections and weighted subtraction reconstructions of the CIRS phantoms with and without SC. We report decupping results using eq.(3.15). As explained in section 3.2.3.2, the choice of polynomial order for interpolating wing data is a crucial part of our SC algorithm. In section 3.3.1 we confirm that

a fourth-order polynomial is optimal. We do this by comparing $S\hat{P}R^{\theta}_{HE}(x,y)$ vs. $SPR^{\theta}_{HE}(x,y)$ using the same methods as used in section 3.3.1, but with $S\hat{P}R^{\theta}_{HE}(x,y)$ assessed at a range of polynomial orders. We display reconstructions vs. polynomial order. Results from the clinical application of SC method are presented in section 3.3.5.

3.3.1. Validation results for CIRS-020 phantom

We report *SPR* and $S\hat{P}R$ results for HE and LE, for various thicknesses and for 3 representative angles: $0°$, $12°$ and $23°$. The values for *SPR* and $S\hat{P}R$ can be compared at all 225 pinhole locations, but we select a subset of 10 representative locations for each angle. The projection shadow is itself elongated at high θ relative to that at $\theta = 0°$ and the plate orientation is θ- dependent. Both effects lead to a shift of the pinhole shadows as θ varies. The diagram in Figure 3.10 illustrates. Figure 3.10(a) shows a $0°$ acquisition illustrating the breast shadow region, pinhole locations (dots) and the 10 selected sampling points. Figure 3.10(b) shows the situation at $\theta = 23°$ with the shadow now slightly elongated, the pinhole shadows slightly shifted relative to $\theta = 0°$, and the 10 sampling points indicated.

The sampling points are numbered with indices 1-10, with points 1-5 corresponding to breast edge locations, 1,9,5 to locations along the chest wall, and 6,7,8 to interior locations. Note that while the positions of the 10 sampling points and the breast shadow itself both vary slightly with θ, the relative dispositions of the 10 index points vary only slightly. Location #10 always corresponds to $(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$.

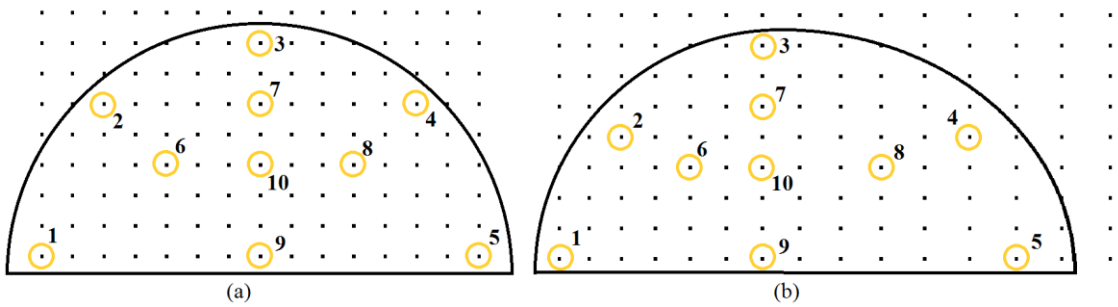

Figure 3.10 Sampling locations for validation. The breast shadow region (solid line) and pinhole shadow (dots) are drawn to scale for a $0°$ acquisition in (a) and a $23°$ acquisition in (b).

A subset of 10 pinhole locations, indicated by circles, is chosen at each angle and validation is done at these locations. Each of the 10 pinhole locations is labeled by an index number.

Table 3.3 displays the HE (left) and LE (right) validation results for various thicknesses and angles. The first column in each sub-table indicates the sampling location via the id number as shown in Figure 3.10. The validations appear as adjacent columns labeled "SPR" and "$\hat{SPR}$". The subscripts (HE or LE), superscript (θ), and argument (x, y) associated with symbols SPR and $\hat{SPR}$ are accounted for by the HE and LE labels on the left and right tables, the angle ($0°$, $12°$ or $23°$) is shown above each SPR vs. $\hat{SPR}$ column pair, and the (x, y) location by the location id in the first column. The error bars on SPR and on $\hat{SPR}$ measurement pairs were far less (much less than 1%) than the difference between the SPR and $\hat{SPR}$ pair, so no standard deviations on SPR or $\hat{SPR}$ are listed. Note that in the 10th row of each table (corresponding to the point with id#10), SPR exactly equals $\hat{SPR}$. This is true by construction since as stated in the 2nd paragraph of section 3.3.1, location #10 always corresponds to $(\bar{x}_c^{\theta}, \bar{y}_c^{\theta})$.

For HE at all thicknesses, the scatter estimates at $0°$ are quite accurate (with SPR and $\hat{SPR}$ differences small.) As θ increases, accuracy drops, especially for sampling points 4 and 5 corresponding to the region where the breast shadow is elongated. For LE comparison of SPR with $\hat{SPR}$, it shows good agreement except at points near the phantom edge. The discrepancy is due to the fact that our simple method of approximating $\hat{s}_{LE}^{\theta}(x, y)$ as a space-invariant constant leads to a discrepancy at edge points, where the true $s_{LE}^{\theta}(x, y)$ shows slow spatial variation(Sechopoulos et al., 2007b, Feng et al., 2014, Boone and Cooper III, 2000).

Table 3.3 Validation results for scatter correction algorithm applied to CIRS-020 phantom.

HE

		0 degrees		12 degrees		23 degrees	
	id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
2 cm	1	23.60%	24.00%	29.30%	27.50%	29.90%	28.70%
	2	29.60%	28.50%	31.00%	29.00%	32.50%	29.10%
	3	30.50%	27.80%	29.20%	28.50%	29.50%	27.40%
	4	29.40%	27.80%	33.80%	32.70%	33.80%	32.90%
	5	25.10%	25.70%	31.10%	33.30%	28.10%	28.00%
	6	34.00%	34.30%	37.50%	38.30%	38.60%	37.60%
	7	36.30%	32.90%	37.00%	34.90%	38.20%	36.60%
	8	35.00%	34.90%	39.50%	39.60%	39.60%	40.40%
	9	35.40%	37.60%	38.40%	41.60%	36.80%	40.80%
	10	37.10%	37.10%	39.50%	39.50%	43.70%	43.70%

		0 degrees		12 degrees		23 degrees	
	id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
3 cm	1	35.80%	36.00%	41.60%	39.00%	44.40%	43.30%
	2	40.20%	39.00%	43.40%	39.50%	43.80%	39.40%
	3	36.80%	35.70%	38.70%	37.30%	39.70%	36.40%
	4	36.60%	36.30%	42.80%	42.20%	45.60%	44.20%
	5	33.00%	34.30%	38.50%	39.70%	38.00%	40.90%
	6	46.40%	46.30%	51.40%	49.20%	48.90%	49.40%
	7	48.10%	45.50%	48.00%	45.60%	49.40%	46.50%
	8	46.40%	45.60%	51.80%	51.70%	51.50%	53.50%
	9	45.00%	48.10%	49.60%	53.50%	45.00%	51.60%
	10	50.30%	50.30%	54.40%	54.40%	54.00%	54.00%

		0 degrees		12 degrees		23 degrees	
	id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
4 cm	1	46.40%	47.90%	46.20%	49.10%	50.90%	49.10%
	2	48.30%	48.90%	53.30%	54.40%	55.20%	50.10%
	3	53.20%	53.10%	52.00%	46.80%	51.80%	44.90%
	4	48.90%	50.40%	46.80%	46.70%	56.40%	55.30%
	5	45.40%	46.00%	48.50%	49.70%	45.10%	48.00%
	6	57.40%	56.00%	61.20%	64.10%	61.60%	69.40%
	7	59.40%	56.90%	65.10%	59.30%	64.60%	60.00%
	8	59.00%	57.10%	59.00%	60.10%	67.70%	66.80%
	9	59.50%	60.80%	64.20%	69.00%	65.40%	72.10%
	10	62.40%	62.40%	66.20%	66.20%	65.40%	65.40%

		0 degrees		12 degrees		22 degrees	
	id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
5 cm	1	58.60%	65.20%	58.70%	58.60%	55.40%	58.50%
	2	65.70%	68.00%	60.70%	58.20%	63.80%	58.20%
	3	67.30%	68.70%	56.80%	57.50%	56.70%	54.10%
	4	65.20%	68.30%	60.90%	63.00%	56.20%	58.60%
	5	57.30%	63.00%	59.10%	65.70%	64.70%	74.10%
	6	70.10%	71.00%	71.80%	68.30%	77.10%	78.00%
	7	75.60%	73.90%	73.60%	69.30%	73.10%	68.20%
	8	73.40%	73.30%	74.70%	74.30%	78.60%	83.80%
	9	69.80%	75.20%	77.00%	80.70%	80.10%	87.90%
	10	77.80%	77.80%	76.70%	76.70%	83.50%	83.50%

		0 degrees		12 degrees		23 degrees	
	id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
6 cm	1	62.70%	67.30%	72.90%	74.00%	83.10%	83.80%
	2	66.90%	71.00%	72.30%	71.80%	67.40%	59.10%
	3	73.50%	76.30%	68.60%	70.60%	66.60%	59.20%
	4	76.00%	75.50%	73.70%	78.20%	70.40%	72.80%
	5	68.40%	71.20%	71.20%	79.00%	68.70%	77.10%
	6	81.70%	80.00%	81.00%	81.80%	93.60%	92.20%
	7	81.70%	79.50%	81.60%	78.90%	84.50%	77.80%
	8	81.20%	79.90%	89.50%	89.20%	94.30%	95.90%
	9	80.70%	81.50%	83.70%	91.70%	88.60%	98.00%
	10	83.30%	83.30%	88.80%	88.80%	95.50%	95.50%

LE

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	25.40%	28.60%	28.70%	26.90%	33.40%	33.10%
2	31.90%	32.30%	33.70%	32.70%	34.40%	34.70%
3	28.30%	31.10%	31.00%	32.10%	38.10%	39.60%
4	28.30%	28.80%	32.70%	35.20%	33.40%	38.50%
5	24.80%	26.20%	31.90%	34.80%	33.30%	43.00%
6	31.60%	29.90%	31.40%	29.30%	31.60%	31.40%
7	30.90%	30.10%	30.30%	29.60%	37.30%	38.50%
8	31.00%	31.10%	36.20%	34.90%	37.80%	38.70%
9	28.50%	30.50%	32.50%	33.50%	29.80%	32.30%
10	32.40%	32.40%	34.50%	34.50%	35.80%	35.80%

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	40.60%	43.20%	42.30%	41.00%	45.10%	45.20%
2	45.60%	45.80%	38.90%	38.40%	43.50%	43.40%
3	41.00%	44.10%	46.60%	49.70%	51.40%	56.80%
4	43.90%	44.40%	45.70%	50.20%	52.40%	61.60%
5	32.30%	33.60%	39.60%	45.80%	49.00%	59.20%
6	40.70%	40.60%	40.20%	35.50%	52.50%	49.10%
7	39.40%	40.30%	44.30%	41.20%	42.90%	45.00%
8	39.70%	39.70%	43.50%	45.00%	55.50%	55.10%
9	44.40%	44.40%	48.00%	48.40%	37.90%	38.20%
10	43.70%	43.70%	41.40%	41.40%	53.10%	53.10%

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	44.90%	48.70%	50.60%	48.00%	42.30%	41.60%
2	41.20%	41.30%	51.60%	48.40%	44.50%	47.20%
3	47.40%	49.10%	55.00%	61.70%	52.90%	58.40%
4	43.80%	43.20%	53.50%	57.60%	47.40%	56.00%
5	46.60%	48.60%	63.20%	75.10%	59.50%	67.30%
6	49.40%	48.60%	51.80%	51.90%	46.60%	46.60%
7	48.00%	48.90%	62.60%	61.80%	55.60%	57.60%
8	50.10%	47.70%	54.30%	52.90%	55.70%	54.10%
9	55.00%	51.60%	57.40%	60.00%	55.70%	55.20%
10	41.00%	41.00%	49.40%	49.40%	52.30%	52.30%

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	64.20%	62.60%	65.60%	62.10%	73.90%	67.80%
2	55.90%	52.20%	66.60%	63.80%	56.50%	51.70%
3	74.60%	73.60%	77.70%	76.00%	79.30%	75.30%
4	46.20%	52.60%	52.90%	63.30%	69.80%	79.40%
5	49.40%	53.90%	76.00%	86.30%	76.90%	90.70%
6	60.90%	58.60%	57.20%	57.20%	72.40%	75.90%
7	53.10%	57.80%	66.60%	72.40%	84.40%	86.80%
8	59.70%	58.70%	62.10%	64.60%	77.30%	78.60%
9	57.60%	60.50%	61.80%	64.30%	75.20%	76.80%
10	58.90%	58.90%	59.30%	59.30%	73.50%	73.50%

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	75.50%	79.40%	71.10%	67.40%	77.10%	69.80%
2	85.00%	83.10%	83.40%	72.50%	69.60%	61.30%
3	81.60%	79.00%	82.80%	80.20%	106.00%	99.00%
4	74.50%	71.70%	80.30%	90.30%	99.70%	127.90%
5	72.30%	77.50%	77.20%	104.50%	91.90%	128.60%
6	67.60%	67.20%	92.10%	85.40%	91.80%	89.20%
7	74.10%	68.60%	70.50%	70.90%	91.10%	91.50%
8	79.00%	80.30%	73.30%	75.90%	94.50%	111.70%
9	75.00%	76.70%	79.20%	82.50%	86.20%	94.30%
10	68.70%	68.70%	79.00%	79.00%	80.80%	80.80%

Table 3.4 summarizes the results of Table 3.3 in the following way: it gives the mean relative error $\bar{r}_{HE}$, defined as

$$\bar{r}_{HE} = \left\langle \left| \frac{\hat{SPR}^{\theta}_{HE}(x,y) - SPR^{\theta}_{HE}(x,y)}{SPR^{\theta}_{HE}(x,y)} \right| \right\rangle_{\theta,id} \tag{3.16}$$

A similar definition applies to $\bar{r}_{LE}$. Both quantities are averaged over the angles and locations of Table 3.3 as indicated by the angle brackets.

Table 3.4 Relative averaged error of SPR and $\hat{SPR}$ for CIRS-020

	$\bar{r}_{HE}$	$\bar{r}_{LE}$
2cm	4.16%	5.21%
3cm	4.11%	4.79%
4cm	4.37%	4.93%
5cm	4.70%	5.98%
6cm	4.25%	8.18%

The validations have used our SC algorithm of section 3.2.3.2 which entails a 4th-order polynomial interpolation of HE wing data. To validate that order 4 is optimal, we computed $\bar{r}_{HE}$ as a function of polynomial order and thickness. The results in Table 3.5 show that orders 4 and 5 are comparable for 2 cm and 4 cm thickness, but that order 4 is far superior (4.3% vs. 22.6%) for the thicker 6 cm phantom. We observed that for low orders (2 and 3), the error is due to underfitting - the quadratic or cubic interpolant does not have the shape flexibility to capture the scatter profile. For higher orders (6, 7 and 8) the error is due to overfitting. Order 5 also suffers overfitting for the thick 6 cm case.

Table 3.5 Relative averaged error as a function of polynomial order. Entries are $\bar{r}_{HE}$ for various polynomial orders and three thicknesses of the CIRS-020 phantom. An entry of X corresponds to scatter interpolants unstable enough to produce negative scatter estimates.

	Order 2	Order 3	Order 4	Order 5	Order 6	Order 7	Order 8
2 cm	13.98%	13.95%	4.16%	3.57%	26.18%	21.73%	38.71%
4 cm	11.98%	12.39%	4.37%	4.83%	8.42%	18.85%	48.02%
6 cm	9.40%	11.30%	4.25%	22.62%	31.07%	X	X

3.3.2. Validation with rounded edge phantom CIRS-011A

We also performed validations with the CIRS-011A phantom, described in section 3.2.1. It lacks iodine inserts, but its rounded edges emulate the shape of a compressed breast. Since the edge structure can affect the penumbral and wing region images used in our SC algorithm, the CIRS-011A serves as a good test. We repeated the validation steps of section 3.3.1 for the CIRS-011A (albeit only for the single thickness of 4.5 cm). We used a 10-point sampling scheme as before with the points placement similar to that shown in Figure 3.10. The results, in a format that follows that of the tables in section 3.3.1, are presented in Table 3.6.

Table 3.6 Validation results for scatter correction algorithm applied to CIRS-011A phantom.

HE

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	46.5%	51.3%	54.6%	55.1%	55.9%	55.2%
2	53.5%	52.8%	56.6%	55.0%	55.9%	50.8%
3	43.8%	41.6%	64.8%	64.7%	67.3%	68.6%
4	52.8%	51.5%	64.9%	71.4%	69.5%	81.0%
5	49.5%	53.2%	60.4%	75.2%	67.0%	86.8%
6	54.2%	53.4%	63.0%	60.6%	63.5%	61.4%
7	54.1%	51.8%	65.1%	66.0%	73.1%	72.9%
8	51.1%	52.4%	69.4%	70.8%	72.7%	75.2%
9	52.5%	53.4%	58.6%	63.7%	65.9%	68.5%
10	52.7%	52.7%	64.0%	64.0%	69.7%	69.7%

LE

	0 degrees		12 degrees		23 degrees	
id	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$	*SPR*	$\hat{SPR}$
1	50.5%	51.9%	53.5%	50.7%	57.5%	54.0%
2	54.8%	54.5%	58.5%	51.9%	51.0%	45.5%
3	56.4%	55.0%	37.5%	35.6%	47.8%	42.2%
4	54.8%	55.0%	46.5%	48.1%	50.4%	51.5%
5	53.9%	53.4%	46.4%	48.8%	48.8%	50.1%
6	65.6%	65.3%	70.7%	69.8%	74.4%	72.5%
7	64.1%	62.8%	67.8%	61.7%	68.0%	59.6%
8	70.7%	69.2%	73.5%	73.2%	73.9%	73.2%
9	67.8%	71.2%	71.0%	76.2%	76.4%	81.3%
10	69.9%	69.9%	75.2%	75.2%	79.3%	79.3%

Table 3.7 Relative averaged error of *SPR* and $\hat{SPR}$ for CIRS-011A

	$\overline{r}_{HE}$	$\overline{r}_{LE}$
4.5cm	4.02%	5.34%

As with the CIRS-020 phantom, the accuracy of the SC algorithm is better for interior points than for edge points. Table 3.7 shows the mean relative errors $\overline{r}_{HE}$ and $\overline{r}_{LE}$ defined in eq.(3.16).

3.3.3 Projection and reconstruction results

The task of the SC is to improve the visibility of the iodine signal in the weighted subtraction reconstruction. In this section, we illustrate the efficacy of the SC algorithm with quantitative and qualitative results using the iodinated CIRS-020 phantom.

Figure 3.11 illustrates the scatter corrections for the iodinated CIRS-020 phantom by displaying projection images acquired under an assortment of conditions. For HE projection images, we display $-\log\left[I_{HE}^{\theta}(x,y)/\tilde{b}_{HE}^{\theta}(x,y)\right]$ as the non-SC version and $-\log\left[\left(I_{HE}^{\theta}(x,y)-\hat{s}_{HE}^{\theta}(x,y)\right)/\tilde{b}_{HE}^{\theta}(x,y)\right]$ for the SC version. The LE versions follow similarly. The HE images suffer from a cupping effect that is mostly corrected by the scatter correction. Scatter correction improves the contrast of the LE images. Note that in the HE images of Figure 3.11, the non-SC and SC images Figure 3.11(a) and (b)) clearly show decupping. However, inspection of the steep ($\theta = 23°$) acquisitions of Figure 3.11(c) and (d) show that decupping uncovers a shading artifact (increasing intensity left to right) in the SC image. This shading is expected and is due to the radiometry of the steep-angle cone-beam acquisition.

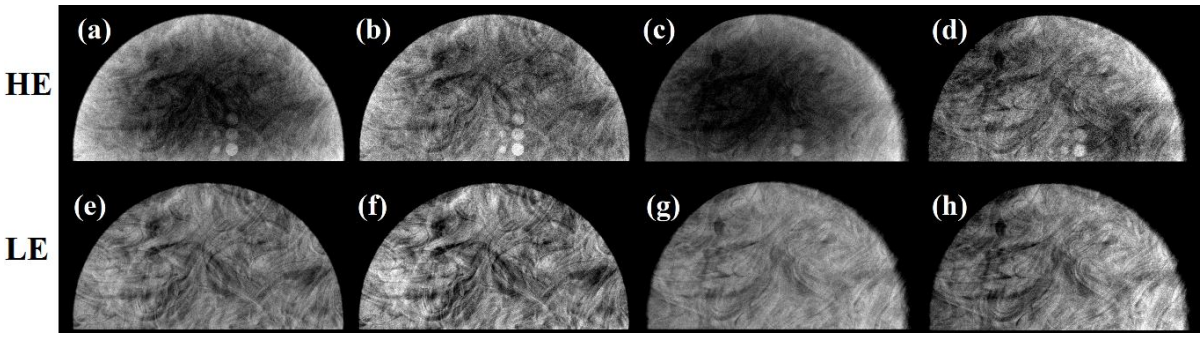

Figure 3.11 Projection images with and without scatter correction for 4cm thick CIRS-020 phantom. Top row: HE images. Bottom row: LE images. Images (a) and (e) at $\theta = 0°$ are non-scatter-corrected, (b) and (f) corrected at $\theta = 0°$, (c) and (g) non-corrected at $\theta = 23°$, and (d) and (h) corrected at $\theta = 23°$. Each SC and no-SC pair uses the same grey scale display window width though the windows are shifted to enhance visual clarity.

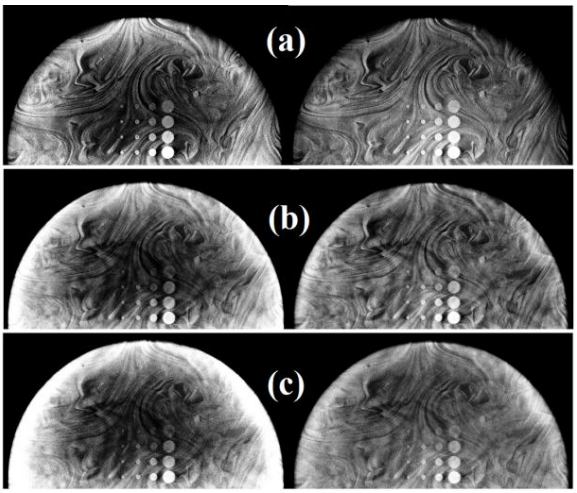

Figure 3.12 Reconstruction of HE data with and without SC. Shown are the reconstructed slices of the CIRS020 phantom at (a) 2cm (b) 4cm and (c) 6cm. The left, right images of each pair are no SC and SC, respectively.

Since scatter effects are most pronounced at high energy and the final reconstruction is a weighted subtraction of HE and LE reconstructions, we separately display HE reconstructions in Figure 3.12. The main effect of the SC is to eliminate the cupping artifacts along the breast periphery. As seen in Figure 3.12, as thickness increases, the cupping artifact becomes more severe.

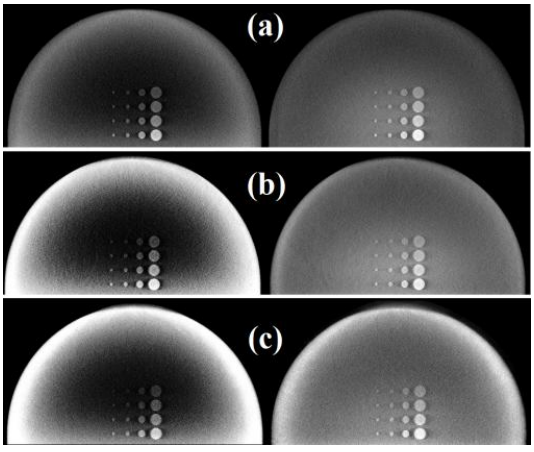

Figure 3.13 Weighted subtraction in the reconstructed domain for CIRS-020 phantom. We use HE and LE reconstruction followed by weighted subtraction to derive the reconstructed slices with iodine inserts. (a), (b) and (c) are for the CIRS-020 phantom at 2, 4 and 6 cm, respectively. The left image of each pair is for no SC and the right image uses SC data.

Figure 3.13 shows a weighted subtraction in the reconstruction domain for the CIRS-020 phantom at 2, 4, and 6 cm. The weighted subtraction reconstruction is designed to suppress the anatomical variation, and comparison of the Figure 3.13 reconstructions to the HE-only reconstructions in Figure 3.12 shows that the "swirl" pattern due to anatomical variability is suppressed. The cupping artifact propagates into the reconstruction, but SC greatly reduces the degree of cupping. For the thicker 4 and 6 cm phantoms the decupping is greatly improved but incomplete and a rim of bright pixels remains.

The SC reconstructions in Figure 3.13 used a 4th order polynomial for interpolating the HE wing data. In Table 3.5 we showed that order 4 was optimal. Figure 3.14 displays weighted subtraction reconstructions of the 4 cm CIRS-020 phantom as a function of polynomial order.

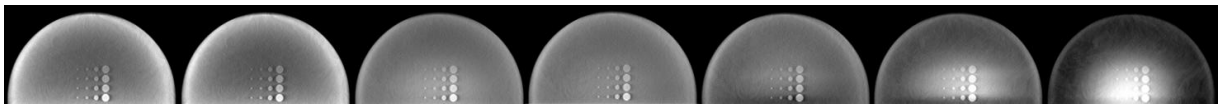

Figure 3.14 Weighted subtraction SC reconstructions using different polynomials orders. The images above are all reconstruction-then-subtraction results for the 4cm CIRS-020 phantom. From left to right, the orders of polynomial used to do wing interpolations are 2, 3, 4, 5, 6, 7 and 8, respectively. The weighting factor and grey-scale windows are identical.

Two important features of SC in the subtracted reconstruction are that: (1) The iodine signal intensity in a given slice becomes independent of the phantom thickness. (2) The decupping due to SC leads to improved iodine signal visibility in a given reconstructed slice.

The first feature is demonstrated in Figure 3.15 which plots iodine signal intensities in the slice containing iodine inserts reconstructed from phantoms of differing thickness. Signal intensities are computed by averaging pixels in signal regions as displayed in Figure 3.3 (b). The pixel averages are taken only up to a radius 0.8 that of the full signal radii in order to avoid edge effects. Different thicknesses yield different amounts of scatter, which, if uncorrected, propagates into the subtracted reconstruction and alters signal values. Our SC does not yield absolute iodine quantitation in the subtracted reconstruction; this goal is intrinsically difficult due to the limited angle nature of DBT reconstruction (Puong et al., 2008). However, by removing

thickness dependence from iodine signal intensity, our SC may support improved iodine quantitation.

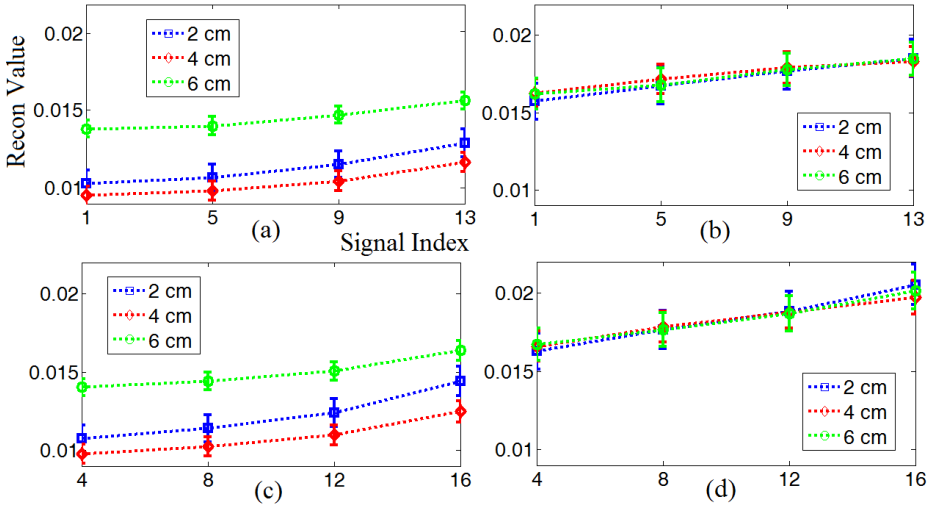

Figure 3.15 Reconstructed intensities of iodine inserts. The ordinate is the reconstructed signal value and the abscissa the index of the signal as displayed in Figure 3.3 (b). (a) No SC. Signals 1, 5, 9, 13 for thickness 2, 4, 6 cm. (b) Same as (a) but with SC. (c) No SC. Signals 4, 8, 12, 16 for thicknesses 2, 4, and 6 cm. (d) Same as (c) but with SC. Note that in (b) and (d), the curves are coincident, illustrating the independence of signal with thickness.

Figure 3.16 shows profiles of the reconstructions of the 4 cm thick phantom of Figure 3.13(b). The profiles clearly show the decupping due to SC in the reconstructed image. In addition, the signal contrast is improved in the SC profiles.

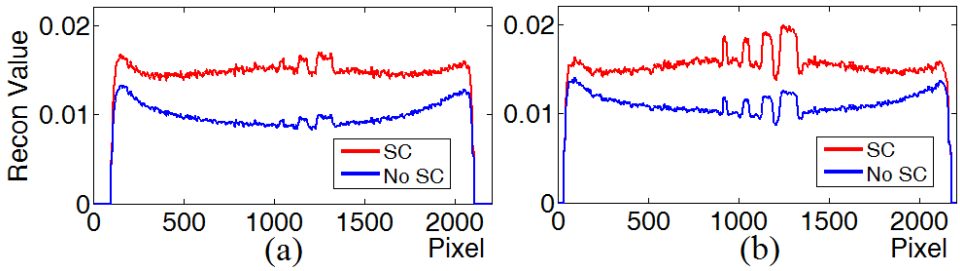

Figure 3.16 Profiles of subtracted reconstruction in signal slice for 4 cm phantom. (a) Profile along centers of top row iodine inserts. (b) Profile along centers of bottom row of iodine inserts.

We evaluated the degree of decupping afforded by SC, using as a metric t_{cup} of eq.(3.15). In eq.(3.15), $\bar{\mu}_{center}$ is the pixel average over a circular region of radius 40 pixels centered on the midline, 4.5 cm from the chest wall. The average $\bar{\mu}_{edge}$ is over 5 circular regions located along the phantom periphery and centered 0.85 cm from the edge. The 5 regions are equiangularly spaced at $45°$, and each of the 5 circular regions is the same size used for $\bar{\mu}_{center}$. None of the 6 regions overlap the positions of the iodine inserts. We averaged t_{cup} over 5 slices including the reconstructed slice in Figure 3.13 ± 2 slices, except for 2cm case where we evaluated 5 slices from the top since the top slice is the signal plane shown in Figure 3.13. For each of the averaged t_{cup} values, we computed a standard deviation using the individual t_{cup} values in each slice. Table 3.8 shows the results for phantoms of thicknesses 2, 4 and 6 cm. As seen in Table 3.8, the degree of decupping is nearly perfect for the 4 cm case, slightly under-decupped for 6 cm and slightly over-decupped for 2 cm.

Table 3.8 Decupping results, t_{cup} for CIRS020 reconstructed images

	2cm	4cm	6cm
No SC	23.0% ±2.2%	24.4% ±2.89%	18.8% ±2.59%
SC	-7.5% ±1.9%	-2.2% ±1.79%	9.3% ±3.01%

In this section we have illustrated the decupping effects of SC on projection data and reconstructions and we demonstrated that SC removes breast thickness dependence from the iodine signal intensity.

3.4 Anecdotal clinical results for scatter correction applied to CE-DE digital mammography

We had an opportunity to apply our SC algorithm to one clinical case for a patient previously diagnosed positive for breast cancer. Our protocol allowed only CE-DE mammography, not DBT. We performed a CE-DE mammographic CC view acquisition using 49 kVp for HE and 29 kVp for LE. We followed a patient-dose protocol that required use of

automatic exposure control. The dose details are reported below. An HE/LE pair (0.3/1.15 mGy) was obtained at time 1 and a second HE/LE pair (0.21/0.61 mGy) at time 2 two minutes later. The contrast material might have changed slightly (diffused or concentrated) during this interval. The image pair at time 1 was obtained using an anti-scatter grid. Importantly, at time 2 the HE image had no grid though the LE image was obtained with a grid. The patient's compressed breast thickness was 5.4 cm.

The fact that no grid was applied for the time 2 HE acquisition offered us an opportunity to test the SC algorithm. We performed a weighted subtraction of the time 2 HE acquisition (with our SC algorithm applied) and the time 2 LE acquisition with scatter correction afforded by a grid. The result is shown in Figure 3.17(c). Figure 3.17(b) shows the same result but with no SC algorithm applied to the HE acquisition. The anecdotal results are stunning, with the tumor clearly demonstrated in Figure 3.17(c) and absent in Figure 3.17(b). Reduction, due to the HE SC algorithm, of the cupping artifact in Figure 3.17(c) helps to reveal the tumor.

We can gain confidence that the results in Figure 3.17(b) and (c) are not artifactual by performing a weighted subtraction on the HE and LE time 1 acquisitions. Since these were obtained with a grid, they are scatter corrected. The result shows a lesion, along the left periphery, which is similar to that seen in Figure 3.17(c). The lesion in Figure 3.17(c) appears to be of higher contrast and slightly different structure compared to that in Figure 3.17(a). However, we do not know if this difference is due to contrast kinetics between time points 1 and 2.

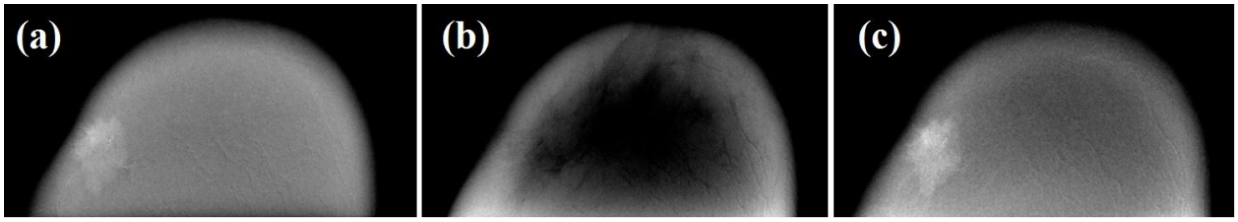

Figure 3.17 Patient study for DE mammography. (a) Weighted subtraction at time 1 with SC for HE and LE provided by the grid (b) Weighted subtraction at time 2 projection data with no SC applied to the HE data and grid SC for the LE data (c) Same as (b) but with our SC algorithm applied to the HE data.

3.5 Discussion and conclusion

We have developed an initial version of our SC algorithm using physical acquisitions with breast phantoms. The use of phantoms allows validation through the use of pinhole array measurements. The algorithm contains site specific quantities $\overline{SPR}^{\theta}_{HE}(\bar{x}^{\theta}_{c}, \bar{y}^{\theta}_{c})$ and $\overline{SPR}^{\theta}_{LE}(\bar{x}^{\theta}_{c}, \bar{y}^{\theta}_{c})$ that must be determined by calibration procedures.

Our studies entailed only CC views. If MLO views become standard for CE-DE-DBT protocols, then the SC algorithm must be adapted to this geometry. For the CC view, the 1-D polynomial interpolation through the wing, penumbral and object regions is done on trajectories along the x-direction, but for the MLO view, trajectories along x include the breast edge on one side and the chest wall on the other. For the MLO view, by tilting the trajectory so that it is at an angle with respect to x, the trajectory then includes wing and penumbral regions on each side of the breast. The 1-D SC interpolation schemes can then be applied, though further work is needed to study differences in the interpolation parameters for MLO vs. CC views.

The SC algorithm needs further improvement in estimating scatter in the penumbral region. Inspection of the SC reconstructions in Figure 3.13 (right column) shows a thin bright ring artifact along the breast periphery that is worse for thicker breasts. The SC algorithm overestimates scatter in the penumbral region as seen in Figure 3.4(d), and this overestimate leads to the artifact.

One might suspect that SC could improve absolute or relative quantitation of iodine concentration in the reconstruction. However, Puong et. al. (Puong et al., 2008) show that iodine quantitation is intrinsically difficult for CE-DE-DBT. The difficulty is due to the severely limited angular acquisition of DBT. However, our results in Figure 3.15 show one useful aspect of SC: the reconstructed iodine voxel values become independent of breast thickness.

In CE-DE-DBT, the visualization of iodinated lesions is improved by anatomical background suppression and K-edge enhancement of the iodine signal. Nevertheless, cupping effects may lead to masking effects for iodine signals located along the breast periphery. Scatter correction may be quite useful in addressing this problem; further work is needed to assess the degree of improvement.

Finally, one might avoid the complexities of algorithmic SC by using an anti-scatter grid. For tomosynthesis, the grid vanes would need to be oriented in the y-direction to avoid angular blocking. However, any such use of grids entails a dose penalty due to absorption of primary photons by the grid.

We presented an algorithm designed to accomplish rapid patient-specific scatter correction for CE-DE-DBT. For a given site, it requires simple calibration measurements obtained from a pinhole SPR measurement of a standardized phantom. We validated its accuracy for breast emulating phantoms. The resulting scatter estimate was showed to reduce cupping artifacts in reconstructions and to remove the effect of phantom thickness on reconstructed values of iodine inserts.

Chapter 4 Impact of Acquisition, Subtraction and Reconstruction Strategies on Dual-Energy Contrast Enhanced Breast Tomosynthesis

Part of (Secs 4.1 to 4.4), this chapter is a paraphrase of an article (Chen et al., 2013) that appeared in the 2013 conference proceedings of the SPIE Medical Imaging Conference. Unlike the work in Chapter 3, which was by and large my own, I share much credit in Secs 4.1 to 4.4 of this chapter with co-authors Lin Chen, Yue-Houng Hu, Wei Zhao and Gene Gindi. In particular, the basic notion of interleaved acquisition is due to Dr. Zhao. Lin Chen (first author) shared the detailed work with me but I feel that my contribution to the detailed work was about 50%. Comments from Dr. Zhao and Dr. Gindi are appreciated as is help from Yue-Houng Hu with acquisitions and comments. As far as Sec 4.5 on OSTR reconstruction, the work is my own and was presented in (Lu et al., 2014b). Because much of the chapter stems from a self-contained paper, certain basic points already introduced in Chapter 1 and 2 are briefly repeated.

4.1 Introduction

Contrast Enhanced Digital Breast Tomosynthesis (CE-DBT) is clinically useful (Gavenonis et al., 2012). It retains the 3D advantages of DBT over conventional digital mammography and allows for enhanced display of structures including masses and vasculature. It does so by cancelling background variations in tissue attenuation so that the iodinated structures are more easily seen. A relevant image quality metric to measure performance is SDNR (signal-difference-to noise ratio). CE-DBT may find use as a screening tool, where the total dose is the same as that of a conventional DBT scan. Alternatively, it could be used as a higher dose diagnostic tool following a suspicious finding from a screening conventional mammographic scan. In this paper we constrain CE-DBT acquisition to a dose equal to that of conventional DBT which is itself equal to that of conventional mammography. Like breast MR, CE-DBT can characterize the contrast agent uptake kinetics of a tumor but it displays anatomy at a higher

spatial resolution than breast MR. In this paper, we are not concerned with contrast kinetics, but with obtaining high SDNR values at a single time point.

In DBT, CE strategies to cancel tissue background and highlight iodinated structures include temporal subtraction (TS) and dual-energy (DE) imaging. Both TS and DE have problems in the propagation of artifacts into the reconstruction, but here we focus on DE-DBT.

In CE-DE-DBT, contrast injection is followed by a high energy (HE) acquisition (above the iodine K-edge) and a low energy (LE) acquisition. As shown in Figure 4.1 (b), CE-DE-DBT can be done by a "two sweep" acquisition protocol. However, there are important advantages in instead doing a one-sweep "interleaved" acquisition (Figure 4.1 (c) (d)) in which a HE acquisition at one angle alternates with a LE acquisition at the next angle while the X-ray tube sweeps through its arc. Advantages of interleave vs. two-sweep acquisitions (Hill et al., 2012) include (1) less artifacts due to patient motion (2) less patient time spent under uncomfortable compression. Due to these advantages, our work focuses on the interleaved acquisition protocols in Figure 4.1 (c) (d). Figure 1 summarizes acquisition and processing of conventional DBT, double-sweep DE-DBT, and two ways of doing interleaved DE-DBT.

4.2 Methods

4.2.1 Acquisition strategies

We presume that for all the acquisition protocols of Figure 4.1, the X-ray tube is swept continuously (not in step-and-shoot mode), and the kVp and filter switching is fast enough for the interleave schemes. The sweep times are all equal, so the double-sweep method of Figure 4.1 (b) takes at least twice as long as the other cases. All acquisitions are presumed to be acquired at a dose equal to that of the screening DBT case of Figure 4.1 (a). In this case, the dose from each interleave exposure at a given angle is twice that of an exposure from the same angle in the double-sweep method. Angles are equispaced by $\Delta\theta$ .For clarity, only 6 angles are shown though our acquisitions will use more. We note that if a step-and-shoot mode is possible, then the HE and LE acquisitions in the interleave scans can be acquired with $\Delta\theta = 0$ and if a photon-

counting detector is used, HE and LE acquisitions could be acquired simultaneously. But here, we exclude these two possibilities.

Conventional data processing for CE-DE-DBT is shown in Figure 4.1 (b) (c). This entails a weighted subtraction of HE and LE projection data followed by reconstruction of the subtracted data. Subtraction, designed to remove anatomical background variability from the CE reconstruction, can create artifacts due to patient motion. The interleave scheme of Figure 4.1 (c) is less susceptible to patient motion artifacts than the double-sweep method. However, this interleave method incurs a subtraction artifact even in the absence of patient motion. Since the angular difference $\Delta\theta$ in the interleaved HE, LE acquisitions creates a shift in the associated image pairs, weighted subtraction of the shifted image pairs creates artifacts that can propagate into the reconstruction and thus lower image quality. We shall refer to the scheme in Figure 4.1 (c) as "subtract/recon" or SR.

We propose to reduce these types of artifacts by the alternate strategy of first reconstructing the HE and LE data separately and then subtracting the reconstructions as shown in Figure 4.1 (d). We call this the RS (reconstruct-then-subtract) strategy. The actual reconstruction algorithm can also affect the propagation of subtraction artifacts, and we explore the effects of FBP and SART on SDNR. The algorithms can be used in an SR context or an RS context.

Our purpose is to explore the effects on lesion SDNR of six acquisition and processing cases. Let DS mean double-sweep and IL mean interleave. As defined previously, RS and SR refer to the reconstruct-then-subtract and subtract-then-reconstruct strategies. The six cases are then: DS-SR-FBP, IL-SR-FBP, IL-RS-FBP, DS-SR-SART, IL-SR-SART and IL-RS-SART.

Note that our focus here is in the effects on lesion SDNR of subtraction artifacts in the presence of background variability under the 6 acquisitions. Many other factors can affect lesion SDNR, such as contrast kinetics during acquisition, spectrum tailoring, detector characteristics and other factors, but our work here does not consider these. Also, we do not model patient motion in this study even though we have pointed out that interleave acquisition is less susceptible to patient motion artifacts than double sweep acquisitions.

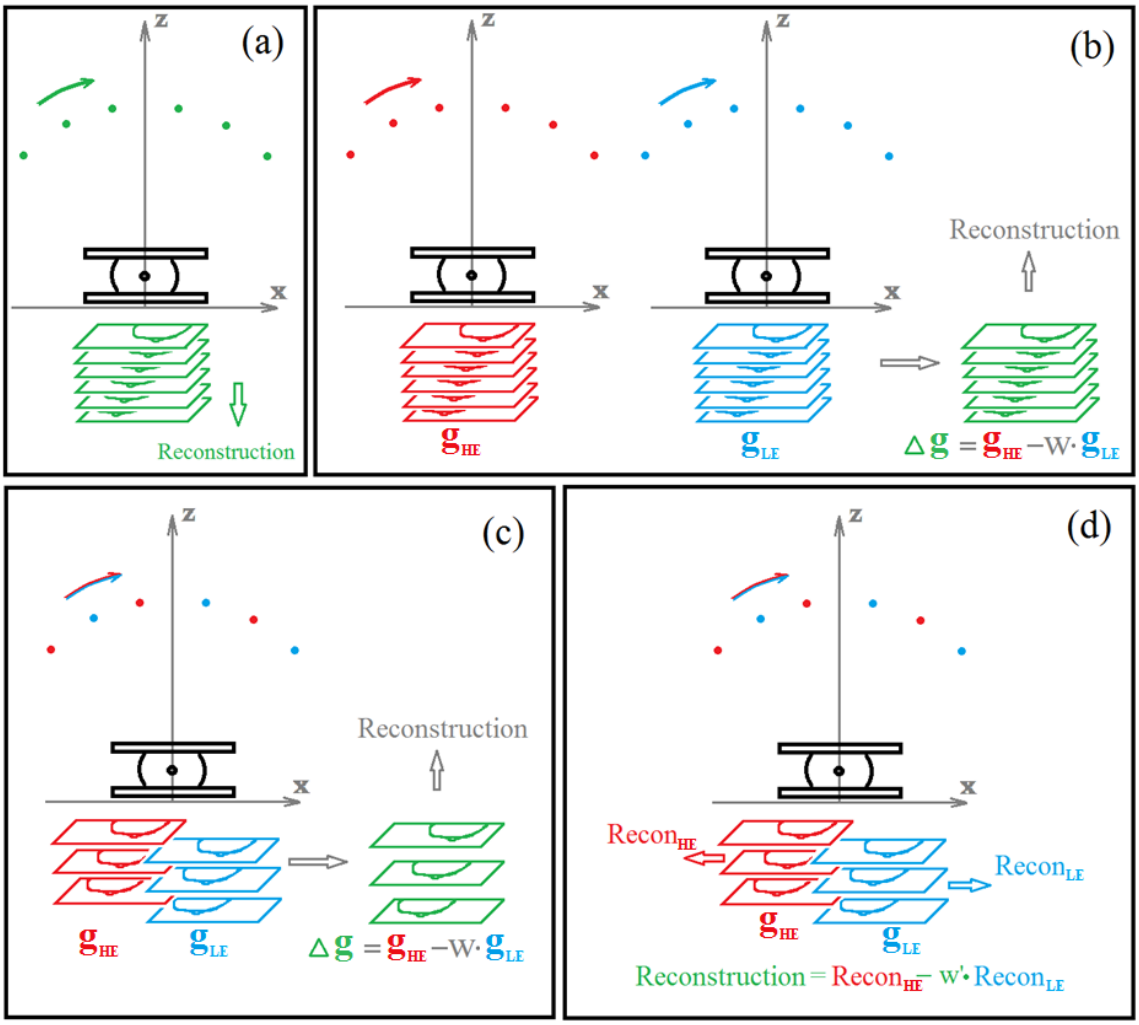

Figure 4.1 Acquisition protocols. (The 3D acquisition and processing takes place in the X-Y-Z space, but for clarity we do not display the Y-axis.) A compressed breast is shown in the X-Z plane and the detector assumed coincident with the X-Y plane. The sketch is not meant to be to scale. Dots indicate source positions. (Discrete dots imply a step-and-shoot acquisition, but in experiments we use continuous tube motion.) For clarity, only 6 source positions are depicted. A green dot is for conventional DBT kVp, red for HE kVp and blue for LE kVp acquisitions. The stacks of rectangles indicate the projection data (after the log step). "Reconstruction" indicates the application of a reconstruction algorithm to the projection data. (a) Conventional DBT (b) Double-sweep DE-DBT. Here the Δg is the weighted subtraction, with weight w, of projection data $\Delta g = g_{HE} - w \cdot g_{LE}$ which is then reconstructed (c) Interleaved CE-DBT with weighted subtraction of adjacent scans followed by reconstruction. The vertical offset between the stacks

of g_{HE} and g_{LE} projections indicates that associated g_{HE} and g_{LE} pairs are obtained from slightly different angles and are therefore laterally shifted in the x-direction. (d) Interleaved CE-DBT with separate reconstruction of HE, LE data denoted by Recon_{HE} and Recon_{LE}. This is followed by weighted subtraction, with weight w', in the reconstruction domain.

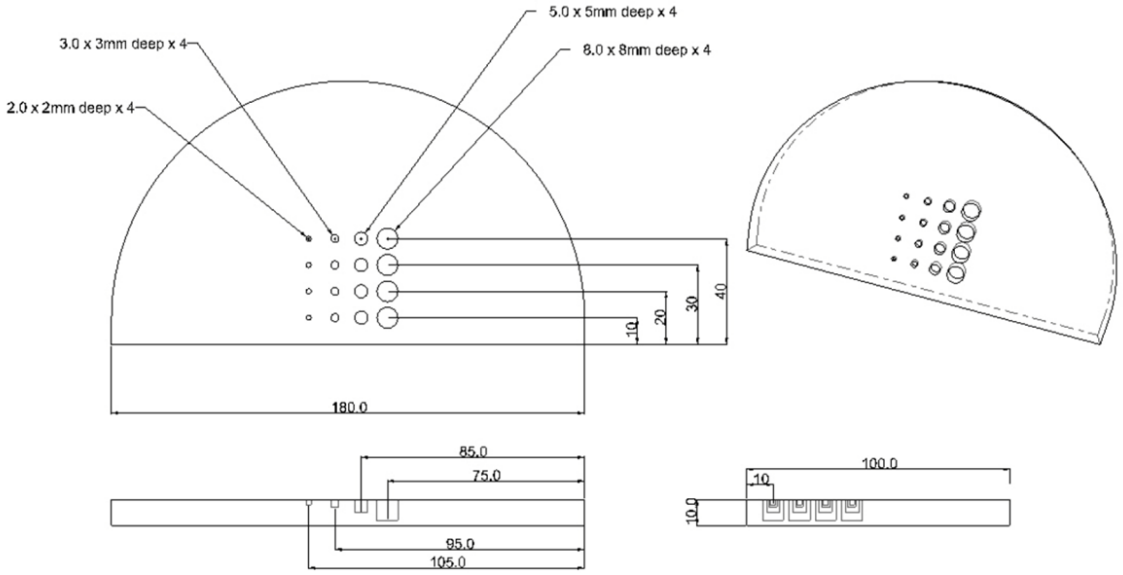

Figure 4.2 Geometry of the iodine inserts. All dimensions are in mm.

4.2.2 Phantom

We acquired phantom images on a prototype Siemens Mammomat Inspiration DBT unit. Figure 3.1 shows the phantom as it is situated in the scanner. The phantom comprised four 1 cm thick semi-circular CIRS model 20 layers designed to mimic the spatial variability of adipose and glandular tissue and the phantom was designed to include 50% adipose and 50% glandular. The 3rd slab (inset in Figure 3.2 (a)) was modified to include a 4 by 4 contrast-detail array of cylinders filled with solid iodine inserts. Figure 4.2 shows the insert geometry. The cylinders in each column are 2, 3, 5 and 8 mm in diameter with the depth of each cylinder equal to its height.

The tops of all 16 cylinders lie in one plane, and it is this reference plane which we will use for in-plane SDNR evaluation. The iodine concentration in each row is (top to bottom) 1, 2, 3 and 5 mg/ml.

4.2.3 Acquisition details

Since our DBT unit could not acquire data in interleave mode, we used double-sweep acquisitions to obtain both DS data (Figure 4.1 b) and IL data (Figure 4.1 c, d). We did this in the following way: To get DS data, we did two sweeps at 71 mAs (H) and 71 mAs (L) and kept all projection data. To get IL data, we did two sweeps at 140 mAs (HE) and 140 mAs (LE) and then culled alternate projections to obtain HE, LE pairs as shown in Figure 4.1 c, d. The resulting interleaved acquisition comprised 12 HE acquisitions and 12 LE acquisitions spanning -25° to $+25^{\circ}$ with an angular separation of $\Delta\theta = 2^{\circ}$.The DS acquisitions comprised 24 HE and 24 LE acquisitions each spanning the same -25° to $+25^{\circ}$ angular range. The tube motion was continuous with an X-ray duty cycle (% of $\Delta\theta = 2^{\circ}$ during which X-rays are generated) of 4.5% for DS and 9% for IL. Because of the small duty cycle the motion blur was negligible. The effective total interleave sweep time was 25 sec and the effective DS sweep time (2 sweeps) was 50 sec. The a-Se detector had a 300 μm thickness and 85 μm detector bin pitch with no bin averaging on readout. For HE, we used a target/filter W/Cu at 49 kVp and for LE W/Rh at 28 kVp. The acquisition leads to about a 50%/50% dose split for HE/LE. The total dose, 1.4mGy for any acquisition, was comparable to that of a conventional DBT screening acquisition. The projection data were corrected for detector gain nonuniformity and for shading due to radiometric effects and the heel effect. No correction of the reconstruction due to scatter was done.

4.2.4 Reconstruction algorithms

For the FBP reconstruction (Mertelmeier et al., 2006) we applied a ramp filter in the X-direction and modified the ramp with a Hamming apodization filter with a cutoff frequency of 0.75 of the Nyquist. A Hamming slice thickness filter to control Z direction aliasing is also

applied in the Z direction with a 0.035 Nyquist cutoff. For a stationary detector geometry, the various filter cutoffs depend on θ as described in (Mertelmeier et al., 2006). We used an ordered-subset version of SART (Wang and Jiang, 2004) with 4 iterations and a constant image as the initial condition. For both FBP and SART, we applied a final median window filtering to the reconstruction, using 5 by 5 pixel window. The rectanguloid reconstruction voxels had dimensions $85\times85\times1000\mu m$. For reconstructions, we used a ray-driven Siddon-based forward projector (Siddon, 1984) and a voxel-driven backprojector based on bin interpolation. The reconstruction programs were implemented using a graphics processing unit (GPU) which resulted in high-performance computing in both FBP and iterative SART reconstruction. The average speedup factor was over 100 relative to a traditional CPU implementation.

4.2.5 Data processing strategies

Let g_{HE} and g_{LE} represent the interleaved projection data after the log step (as in Figure 4.1 c, d). Let $\hat{\mu}$ denote a reconstructed image. Let O denote a reconstruction operator (SART or FBP). Then the SR strategy of Figure 4.1 (c) can be summarized as $\Delta g = g_{HE} - w \cdot g_{LE}$ followed by $\hat{\mu}_{SR} = O(\Delta g)$ The weight $w = 0.3$ was determined using methods in (Hu and Zhao, 2012). For the RS strategy, we first perform two reconstructions $\hat{\mu}_{HE} = O(g_{HE})$ and $\hat{\mu}_{LE} = O(g_{LE})$. Then $\hat{\mu}_{SR} = \hat{\mu}_{HE} - w'\hat{\mu}_{LE}$ completes the RS strategy. We used $w' = 0.2$ derived empirically. In (Gavenonis et al., 2012) a form of RS was used with $O =$ simple backprojection and a post-filter applied to the result. For the DS data processing (Figure 4.1 b), g_{HE} and g_{LE} denote the double-sweep (non-interleave) projection data. Again $\Delta g = g_{HE} - w \cdot g_{LE}$ with w = 0.3 and $\hat{\mu}_{SR} = O(\Delta g)$.

The RS and SR strategies are characterized by an important difference. For the reconstructions in RS, each of the HE and LE detector values are backprojected along the directions from which they were acquired. But for SR, a given detector bin contains contributions from an HE ray and an LE ray each separated by $\Delta\theta$. Therefore, there is an ambiguity in the proper direction to be used for backprojection. One could backproject along the HE ray or the associated LE ray of the interleaved pair, or at some average angle. In our reconstructions we backprojected along the HE direction. The subtraction artifact and ambiguity of backprojection angle associated with SR are not present in RS. This results in RS

reconstructions with far fewer artifacts than SR reconstructions. We will demonstrate this with experimental data below.

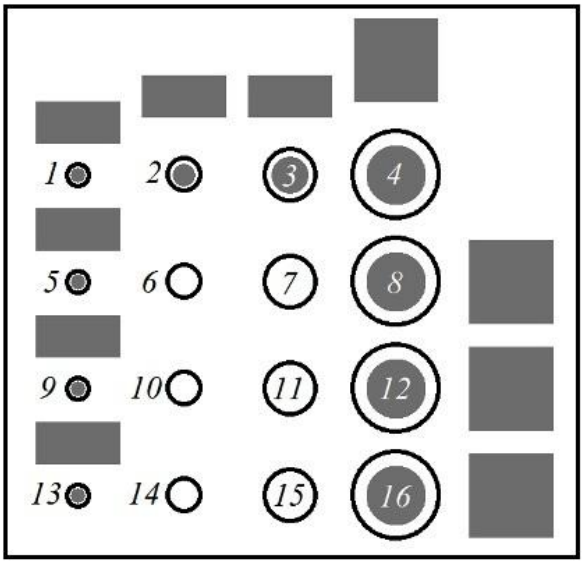

Figure 4.3 The numbers in each circle index the signal. The circular grey regions (70% of signal radius) are signal templates used in SDNR and the adjacent square and rectangular grey regions are the associated background regions used in SDNR calculation.

4.2.6 SDNR calculation

SDNR was calculated for each of the 6 cases mentioned in Section 4.2.1 and for 10 of the 16 signals of the 4 by 4 contrast-detail phantom. SDNR was evaluated in a reference plane containing the tops of all cylinders as seen in the lower left of Figure 4.2. SDNR was defined the (mean signal - mean background) / (standard deviation of background) in a manner similar to (Wu et al., 2009) and (Wu et al., 2004). Note that fluctuations due to the propagation of subtraction artifacts are an important contributor to the means and standard deviation in the SDNR definition. Figure 4 shows the circular signal templates and square background templates used in SDNR calculations. Signals 6, 7, 10, 11, 14 and 15 were not evaluated due to lack of space for fitting a background template. Signals 1, 5, 9 and 13 were evaluated, but we note that the very small diameter of these four signals led to less reliable SDNR values.

Table 4.1 SDNR for each of the 6 acquisition-reconstruction combinations. The signals corresponding to each signal index are shown in Figure 4.3. DS = double sweep and IL = interleave

Signal index	DS-SR-FBP	IL-SR-FBP	IL-RS-FBP	DS-SR-SART	IL-SR-SART	IL-RS-SART
1	1.159	1.641	1.187	1.491	2.682	2.747
2	1.165	-0.188	0.583	1.388	-0.017	1.697
3	0.855	0.658	0.885	1.407	1.323	2.177
4	1.198	0.545	0.971	2.014	1.313	2.247
5	2.420	2.572	2.363	2.388	3.088	4.207
8	1.656	0.604	1.173	2.474	1.656	3.135
9	2.666	2.384	2.752	3.427	3.788	5.399
12	2.144	0.911	1.992	2.542	1.528	4.233
13	3.318	1.581	3.667	3.848	2.289	7.021
16	2.584	1.187	3.118	3.280	2.147	5.358

4.3 Results and Conclusions

Figure 4.4 shows reconstructions in the reference plane. Note that the subtraction artifacts, striations in the interleaved reconstructions indicating incomplete background removal, are more apparent for the SR cases in Figure 4.4 (b) (f) than in the RS cases of Figure 4.4 (c) (g).

Table 4.1 summarizes quantitative SDNR results. One observation illustrated in the table and figure is the following: (1) For our proposed interleave protocol, and for both FBP and SART, the RS method yields superior results relative to the SR method. A second observation is that (2) for projection data collected under any protocol, the SART reconstruction yields better SDNR than FBP. A third observation is that (3) in comparing any DS result to any IL result, there is no consistent superiority of one method vs. the other. However, the particular case of IL-RS-SART is far superior to any DS result. Note that the SDNRs for signal 2 in the IL-SR-FBP

and IL-SR-SART cases are negative. This is due to the fact that signal 2 is of small diameter with a small averaging template, so very few pixels are used to calculate signal difference. In this case, the few noisy pixels in the signal-absent case were of higher intensity on average than those of the signal-present case. Here "noise" is mainly to the propagation of subtraction artifacts, not quantum noise.

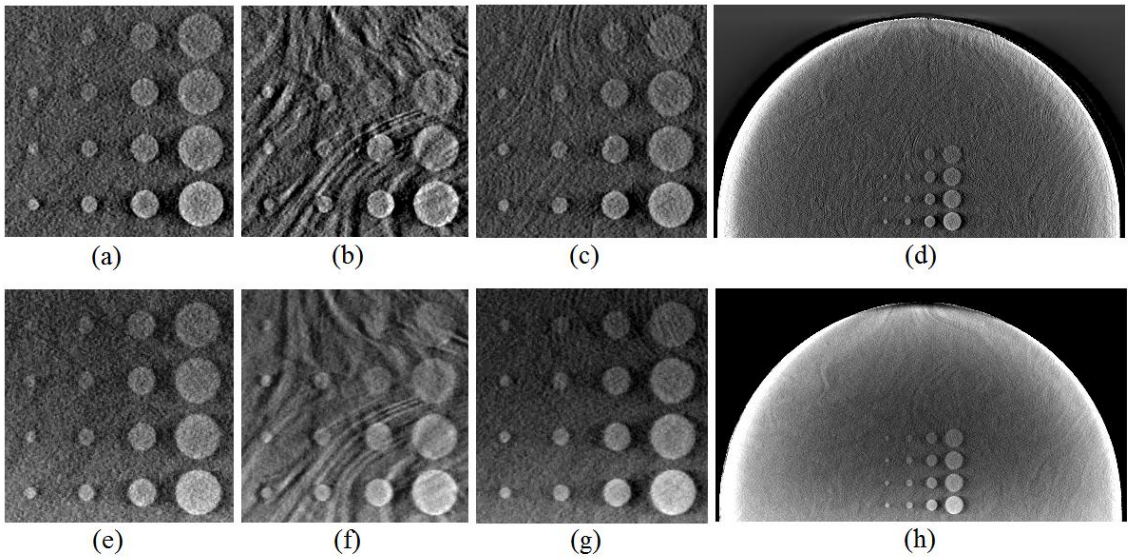

Figure 4.4 Reconstructions in the reference plane. (a) DS-SR-FBP. (b) IL- SR-FBP. (c) IL-RS-FBP (d) IL-RS-FBP. (e) DS-SR-SART. (f) IL-SR-SART (g) IL-RS-SART (h) IL-RS-SART. (DS=double sweep, IL=interleave, SR=subtract-then-reconstruct, RS=reconstruct-then-subtract) The grey scales for each image here are chosen for viewing convenience, but quantitative image quality is summarized in Table 1. The reconstructions show only the signal region, except for (d) and (h) which show the full phantom slice corresponding to (c) and (g).

There are strong practical reasons - less susceptibility to patient motion and less patient time spent under compression - to use interleaved acquisition rather than double-sweep acquisition for DE-DBT, but for continuous tube motion, interleaving incurs unavoidable artifacts in the reconstruction even in the absence of patient motion. For interleaved acquisition, reconstruction followed by subtraction (RS) is preferred (in the sense of SDNR) over subtraction followed by

reconstruction (SR) for DE-DBT, though the choice of algorithm (SART vs. FBP) has an even stronger effect.

4.4 A statistical reconstruction for CE-DE-DBT

4.4.1 Motivation

In the results in Sec.4.3 we observed that the algorithm choice had a strong influence on the image quality, with SART outperforming FBP. This motivated later work (Lu et al., 2014b) in which I tried to apply a statistical reconstruction method to the same problem that was reported in Sec.4.3.

We already posed reconstruction as maximum-likelihood problem in Sec. 2.5 the ML technique, unlike SART, contains a noise model, while SART itself is more realistic than FBP. The Poisson noise model for DBT, discussed in Sec. 2.5, is not accurate in that it ignores detector and readout effects. The DBT reconstruction problem, using *any* method still suffers from highly limited angle acquisition. Nevertheless, I though it worth it to apply statistical reconstruction to CE-DE-DBT.

Further motivation came from the fact that for ordinary (anatomical) DBT, statistical reconstruction techniques have been used with some success when visualizing calcifications (Wu et al., 2004, Das et al., 2011, Sidky et al., 2009, Zhang et al., 2006). Here, we have a different problem: visualizing small iodinated areas using the HE and LE acquisitions.

Statistical reconstruction can also serve as a framework for incorporating regularization (prior information), though regularization can also be incorporated into non-statistical techniques (Lu et al., 2010). However, we can regularize the ML reconstruction to obtain a penalized -ML algorithm where the penalty (loosely speaking the "prior" in a maximum a posteriori approach) is sophisticated: it smoothes the recon (we shall use the shorthand "recon" for "reconstruction") when the fluctuations are small but suspends smoothing (preserves edges) when fluctuations are big. This is tantamount to deciding that low-amplitude edges are due to noise and high amplitude edges due to signal.

I will formulate the penalized likelihood recon as the maximization of an objective function. The choice of algorithm for maximization is important in that it must have desirable properties

(discussed below) and be first enough for clinical use. We adapted an algorithm OSTR (Ordered Subset Transmission Reconstruction) for use in CE-DE-DBT. OSTR was first proposed by (Erdogan and Fessler, 1999) for reconstructing noisy transmission data obtained from an external radioactive source, for purposes of attenuation correction in PET.

Our experimental trials will follow those of Sec.4.3, using the CIRS-020 phantom, but in this work, conducted some time after that of Sec.4.3, an important change was introduced: we will make use of the scatter correction algorithm of Chapter 3 to correct the data (unlike the work in Sec.4.3). We will use an RS reconstruction technique and compare FBP SART and OSTR. All acquisition details for obtaining scanner data are the same as that reported in Sec.4.3. The work in this section appeared as a conference paper in (Lu et al., 2014b).

4.4.2 OSTR reconstruction method

We begin by rewriting eq.(2.9) for the HE and LE cases:

$$\mathbf{I}^{HE} = \mathbf{b}^{HE}\mathbf{e}^{-\mathbf{A}\boldsymbol{\mu}^{HE}} + \mathbf{s}^{HE} + \mathbf{n}^{HE} \tag{4.1}$$

$$\mathbf{I}^{LE} = \mathbf{b}^{LE}\mathbf{e}^{-\mathbf{A}\boldsymbol{\mu}^{LE}} + \mathbf{s}^{LE} + \mathbf{n}^{LE} \tag{4.2}$$

where we have used vector notation, dropping the *m* subscripts. Furthermore, it will be convenient to place HE, LE as a superscript in eq.(4.1) and (4.2) instead of a subscript used previously.

We take $\mathbf{n}^{HE}$ and $\mathbf{n}^{LE}$ as Poisson though, as mentioned earlier, the Poisson model is not entirely accurate. While the Poisson noise model is inaccurate, it does attribute more noise to higher detector readings consistent with any form of radiation measurement. The scatter terms $\mathbf{s}^{HE}$ and $\mathbf{s}^{LE}$ will be known from our SC method of Chapter 3. For FBP and SART, we subtract the scatter terms before reconstruction. For OSTR, we retain the known $\mathbf{s}^{HE}$ and $\mathbf{s}^{LE}$ terms in the objective function to reserve the Poisson nature of the noise. That is, if the noise were exactly Poisson as it was in (Erdogan and Fessler, 1999) for the PET attenuation correction problem, then subtracting the scatter would make the raw data $\mathbf{I}$ no longer Poisson. (Indeed, by keeping $\mathbf{s}^{HE}$ and $\mathbf{s}^{LE}$ in the objective function, we will get a slightly lower variance than if we had subtracted it.) To sum up, the data is assumed independent Poisson with means

$$\bar{\mathbf{I}}^{HE} = \mathbf{b}^{HE}\mathbf{e}^{-\mathbf{A}\boldsymbol{\mu}^{HE}} + \mathbf{s}^{HE} \tag{4.3}$$

$$\bar{\mathbf{I}}^{LE} = \mathbf{b}^{LE}\mathbf{e}^{-\mathbf{A}\boldsymbol{\mu}^{LE}} + \mathbf{s}^{LE} \tag{4.4}$$

The objectives functions that we need to minimize are

$$\begin{aligned}\Phi^{HE}(\boldsymbol{\mu}^{HE}) = &\sum_m \left[y_m^{HE} \log(b_m^{HE} e^{-[A\boldsymbol{\mu}^{HE}]_m} + s_m^{HE}) - (b_m^{HE} e^{-[A\boldsymbol{\mu}^{HE}]_m} + s_m^{HE}) \right] \\ &-\beta \sum_n \sum_{k \in \mathrm{N}(n)} \tfrac{1}{2} \omega_{n,k} \Psi(\mu_n^{HE} - \mu_k^{HE})\end{aligned} \tag{4.5}$$

$$\begin{aligned}\Phi^{LE}(\boldsymbol{\mu}^{LE}) = &\sum_m \left[I_m^{LE} \log(b_m^{LE} e^{-[A\boldsymbol{\mu}^{LE}]_m} + s_m^{LE}) - (b_m^{LE} e^{-[A\boldsymbol{\mu}^{LE}]_m} + s_m^{LE}) \right] \\ &-\beta \sum_n \sum_{k \in \mathrm{N}(n)} \tfrac{1}{2} \omega_{n,k} \Psi(\mu_n^{LE} - \mu_k^{LE})\end{aligned} \tag{4.6}$$

The first terms in eq.(4.5) and (4.6) are simply the log likelihoods $p(\mathbf{I}^{HE} \mid \boldsymbol{\mu}^{HE})$ and $p(\mathbf{I}^{LE} \mid \boldsymbol{\mu}^{LE})$ with means $\bar{\mathbf{I}}^{HE}$, $\bar{\mathbf{I}}^{LE}$ given by eq.(4.3) and (4.4). Constant terms have been dropped from the log likelihood. The second terms in eq.(4.5) and (4.6) are the penalties (priors), which we discuss in detail below. Given eq.(4.5) and (4.6), the HE and LE reconstructions are

$$\hat{\boldsymbol{\mu}}^{HE} = \arg\max_{\boldsymbol{\mu}^{HE}} \Phi^{HE}(\boldsymbol{\mu}^{HE}) \tag{4.7}$$

$$\hat{\boldsymbol{\mu}}^{LE} = \arg\max_{\boldsymbol{\mu}^{LE}} \Phi^{LE}(\boldsymbol{\mu}^{LE}) \tag{4.8}$$

where the maximization is carried out by an iterative algorithm to be discussed. Given eq.(4.7) and (4.8) the final reconstruction using the RS strategy is

$$\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}^{HE} - w \cdot \hat{\boldsymbol{\mu}}^{LE} \tag{4.9}$$

Some questions are (i) what is the algorithm? (ii) what is the regularizer? (iii) what is the weight w?

In the penalty term in eq.(4.5), Ψ is a "potential function" that penalizes differences of neighboring pixel values. For example, take $\Psi(z) = z^2$. Then $\Psi(\mu_n^{HE} - \mu_k^{HE}) = (\mu_n^{HE} - \mu_k^{HE})^2$. In the summation $k \in \mathrm{N}(n)$ means that k is a neighboring voxel of voxel n. In 3-D, this could be the 26 nearest neighbors as shown in Figure 4.5. For our work, we used the 26 nearest neighbors. The $\omega_{n,k}$ weights nearer neighbors more strongly, and in our case, we take

$\omega_{n,k} = (\text{distance from } n \text{ to } k)^{-1}$. As voxel values come closer to one another, Ψ drops in value, thus increasing Φ^{HE}. So the penalty term "encourages" smooth solutions. The scalar $\beta > 0$, weights the overall influence of the penalty versus that of the log likelihood is often chosen by hand or by using some task-performance measure. The likelihood forces data agreement of $\hat{\boldsymbol{\mu}}^{HE}$ with the known measurements $\mathbf{I}^{HE}$, while the penalty forces consistency with some belief like smoothness. For the LE case, everything follows as in the HE case, though the value of β and the parameters associated with Ψ can differ.

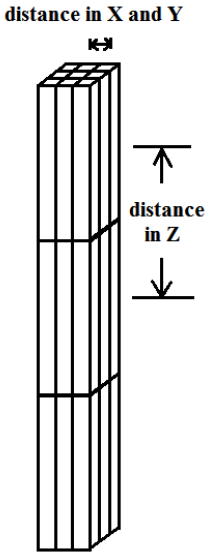

Figure 4.5 Rectanguloid 26 nearest neighborhoods voxels. This is the 26 the nearest-neighbor the neighborhood used in our OSTR reconstruction. The interactions of pixels in the smoothing term are those in the neighborhood.

For a quadratic potential function, the change at one iteration in μ_n is proportional to the difference between μ_n and neighbor μ_k. To see this consider $\Psi(\Delta\mu)$ where we use $\Delta\mu$ to express the difference term, in this case $\mu_n - \mu_k$. An iterative step will cause a change in μ_n that goes as $\frac{\partial\Psi(\Delta\mu)}{\mu_n} \propto \Delta\mu$. Consider a different Ψ that looks like Figure 4.6. For pixel differences $\Delta\mu < \delta$, the penalty is quadratic and acts as before. For $\Delta\mu > \delta$, Ψ is a constant and *no smoothing is applied*. So if an edge is of height $> \delta$, the penalty is "turned off" and the edge is

preserved. For example, at the edge of a calcification or iodine bolus, one might want to suspend smoothing by setting δ appropriately.

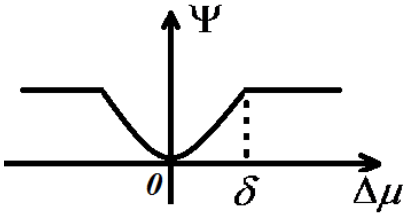

Figure 4.6 Broken parabola potential function.

One problem with using a potential as in Figure 4.6 is that it is non-convex and non-differentiable which leads to optimization problems. Many other "edge preserving" potentials Ψ that are convex and differentiable have been proposed (Lange, 1990, Piyavisetpat et al., 2005, Henschke, 2000) by us and others. The one I used in the OSTR work is the "Lange prior" proposed in (Lange, 1990). It takes the form

$$\Psi(\mu_n - \mu_k) = \delta^2 \cdot \left[\left| \frac{\mu_n - \mu_k}{\delta} \right| - \log\left(1 + \left| \frac{\mu_n - \mu_k}{\delta} \right| \right) \right] \tag{4.10}$$

Figure 4.7 shows the Lange prior compared to a simple quadratic. The user selection of δ controls the distinction between noise and object edges. The sub-parabolic Lange prior is convex and differentiable and preserves edges.

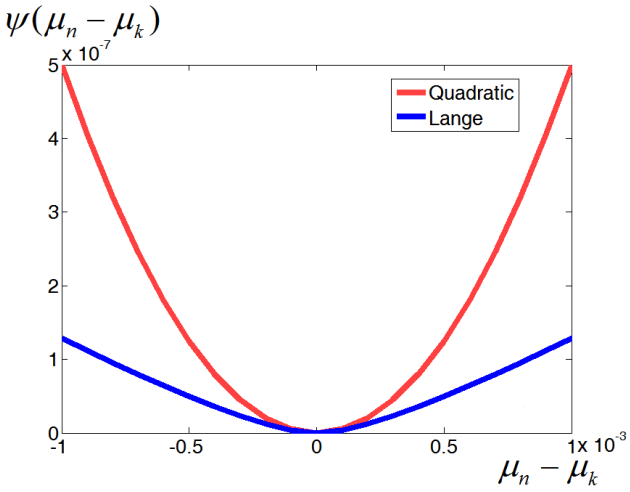

Figure 4.7 Potential functions for different regularizers. Here, $\delta = 0.0002$ for the edge preserving Lange regularizer.

The maximization algorithm for OSTR is complex and is described in detail in (Fessler, 2000). It is parallelizable (all voxels updated simultaneously) and preserves positivity of the reconstructed voxel $\hat{\mu}_n$. OSTR is characterized by the following strategy: At the location of the current iterate on the "surface" Φ, the objective Φ is replaced by a surrogate function - an easier to maximize parabolic objective that satisfies certain properties. Optimizing the parabola guarantees a monotonic increase in Φ. At the new iteration, the process is repeated with new surrogates computed. Since the surrogate is computed for each voxel (parallelizable) and since the surrogates are parabolas, the technique is referred to as "separable paraboloidal surrogates".

The "OS" in OSTR refers to ordered subsets. The OS scheme is used to attain very rapid convergence. In medical imaging, it was first proposed for use in SPECT in (Hudson and Larkin, 1994). The basic idea is as follows: voxels μ_n are updated such they are consistent with the detector readings I_m as well as the penalty. In OS, we introduce a subiteration in which the voxel updates are made consistent (via the likelihood term) with a *subset* of detector readings. After updating voxels with respect to one subset, we repeat for the next subset till all subsets are done. We then move to the next outer loop iteration.

I experimented with subset schemes and used the following: we acquire at 25 angles. All detector readings from every 5th acquisition comprise a subset. Therefore, the 2-D detector outputs from angles 1, 6, 11, 16, 21 form the first subset, those from 2, 7, 12, 17, 22 the second subset, and so on. Therefore we have five subsets. The speedup afforded by OS is considerable. With 5 subsets, we attain a good reconstruction in 30 iterations taking 20 minutes on our CPU/GPU hardware. The same dataset using a single subset (i.e. no subsets) takes 150 iterations and 100 minutes. The use of OS also reduces requirements on GPU memory size considerably.

We've considered the algorithm and penalty, but still need to compute the weight *w* of eq.(4.9) for the RS method. Consider 2 adjacent voxels as shown in Figure 4.8, with the "A" and "G" subscripts indicating Adipose and Glandular. The voxel attenuations are displayed for HE and LE. If we perform weighted subtraction, then the resultant is $\Delta\mu_G$, $\Delta\mu_A$ as shown in the figure. We demand the $\Delta\mu_G = \Delta\mu_A$ in order to suppress the anatomical background. Thus $\mu_G^{HE} - w\cdot\mu_G^{LE} = \mu_A^{HE} - w\cdot\mu_A^{LE}$. Solving for *w* we obtain $w = \frac{\mu_{HE}^G - \mu_{HE}^A}{\mu_{LE}^G - \mu_{LE}^A}$ which is the same as eq.(2.19) which was derived for the SR method. Note also that while in Figure 4.8 the voxels are adjacent, the argument applies to any two voxels. Thus for the RS method, we can use the same weight as in the SR method. We do not need the reconstruction to be linear in order for this to hold.

The results that $w = \frac{\mu_{HE}^G - \mu_{HE}^A}{\mu_{LE}^G - \mu_{LE}^A}$ presumes that the reconstructed $\hat{\mu}$s (the ones in Figure 4.8) are reasonably accurate. However, due to limited angle, noise and other effects, the $\hat{\mu}$ values may be quantitatively inaccurate. If the inaccuracy is severe, then the argument for *w* breaks down. Further investigation is needed. In the next section, we show some results illustrating the OSTR recons using *w* from eq.(2.19) and the results are fairly good.

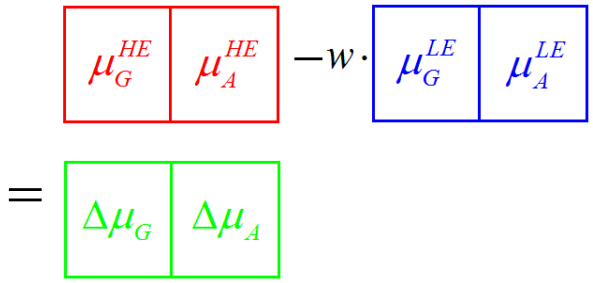

Figure 4.8 Adjacent voxels of glandular (G) and adipose (A) composition and at HE and LE. Weighted subtraction is depicted. The figure is used in the text to explain the derivation of weight *w* for the RS method.

In Sec. 4.2.5 we used w=0.2 for RS and w=0.3 for SR which seems inconsistent with our above argument. However, the lack of scatter correction for the results in Sec. 4.2.5 resulted in inaccurate reconstructions and w=0.2 was chosen empirically. With SC, we use w=0.3 for RS.

4.4.2 OSTR reconstruction method

We used the CIRS-020 phantom and acquisition parameters from Sec. 4.2.2 to study OSTR RS reconstructions and compare these to FBP and SART RS recons.

To see the effects of using quadratic vs. edge-preserving regularizers, we illustrate the OSTR reconstructions of the iodine inserts. In Figure 4.9, the reconstruction with the edge preserving regularizer has sharper edges as expected. (The actual iodine insert physically has a sharp edge.)

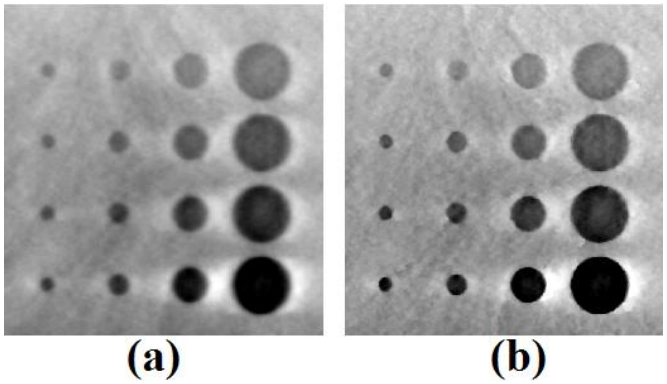

Figure 4.9 OSTR reconstruction of iodine inserts with (a) quadratic smoothing and (b) edge preserving regularizer. Reconstruction parameters: (a) HE recon: $\beta = 250,000$, 50 iterations; LE recon: $\beta = 30,000$, 50 iterations. (b) HE recon: $\beta = 250,000$ $\delta = 0.0002$, 100 iterations; LE recon: $\beta = 30,000$ $\delta = 0.003$, 100 iterations.

To see the effects of different algorithms, we display in Figure 4.10 the reconstruction of the iodine insert region using a highly engineered version of OS-SART versus OSTR with the Lange regularizer. The FBP and SART reconstructions include a post-processing median filter operation that reduces high frequency noise without otherwise affecting image quality.

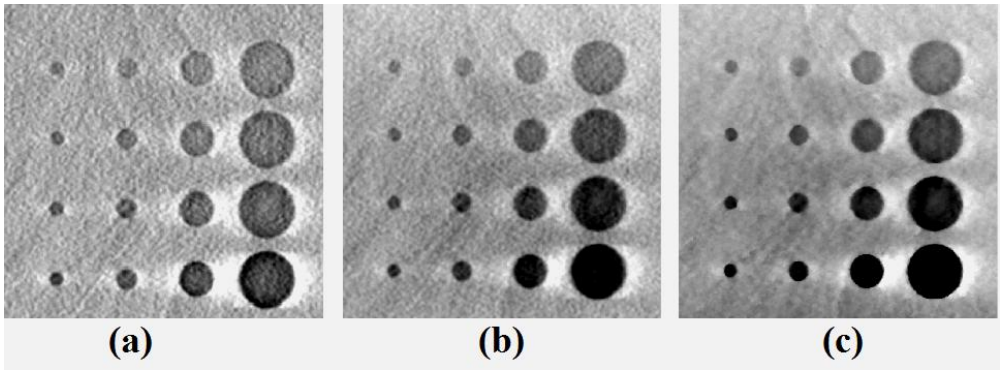

Figure 4.10 Reconstruction of iodine region with (a) FBP (b) SART (c) OSTR. Reconstruction parameters: (b) HE recon: $\lambda = 0.2$, 3 iterations; LE recon: $\lambda = 0.2$, 4 iterations. (c) HE recon: $\beta = 250,000$ $\delta = 0.0002$, 100 iterations; LE recon: $\beta = 30,000$ $\delta = 0.003$, 100 iterations. Note that (a) and (b) were post-filtered by a 5x5 median filter.

The results so far do not appear dramatically different.

It is difficult to appreciate the effects of algorithm unless we look at reconstructions of the smallest and weakest iodine inserts in Figure 4.11. (The one in the upper left corner (I) of the 4x4 array and the one right below the one in the corner (II)). Figure 4.11 shows the results. Anecdotally, OSTR > SART > FBP in terms of fidelity to the true insert. The results in Figure 4.11 are anecdotal and parameters chosen to make FBP SART and OSTR appealing. More systematic study is needed for future work.

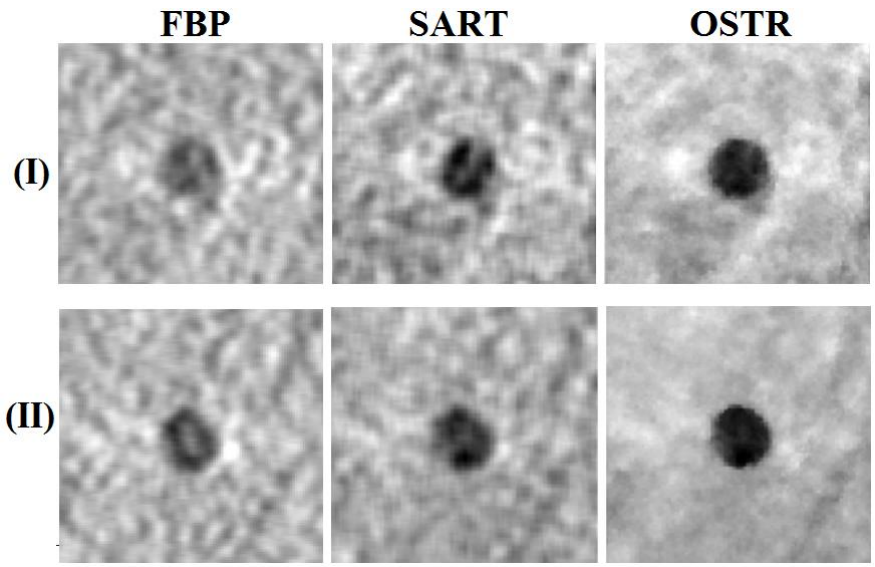

Figure 4.11 Reconstructions of the two smallest and weakest iodine inserts using FBP, SART and OSTR.

We can obtain a quantitative comparison using the SDNR (signal difference to noise ratio) to measure iodine visibility. If "sig" is the signal region and "bckgnd" a background region in the vicinity of the signal, then we use the following definition for SDNR

$$SDNR = \frac{\overline{sig} - \overline{bkgnd}}{0.5\left[\mathrm{STD}(sig) + \mathrm{STD}(bkgnd)\right]}$$

where the bar indicates a pixel average, and STD means "pixel standard deviation" of pixels in the region. Actually for each signal region, we use four background regions in the vicinity of

each of the signal regions as shown in Figure 4.12. The signal region is slightly smaller than the signal itself to avoid edge artifacts. The SDNR averaged over all 16 signals is shown in Table 4.2. In terms of SDNR, OSTR>SART>FBP.

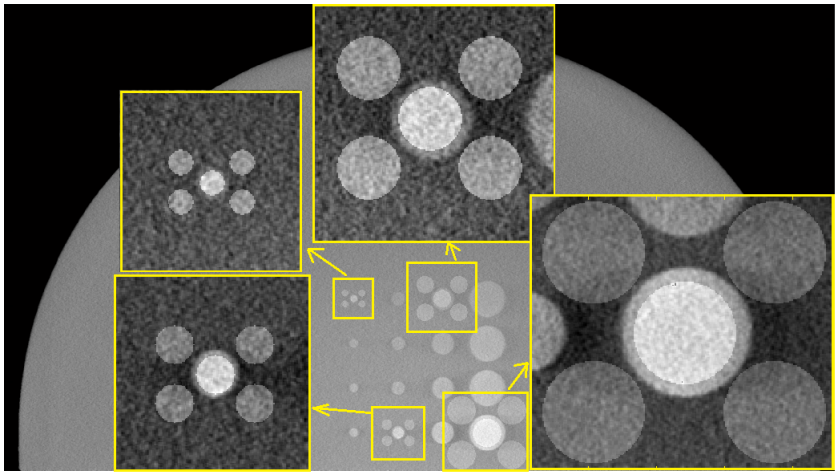

Figure 4.12 Signal and background regions used in SDNR calculation. Each signal region is associated with four nearby background regions.

Table 4.2 SDNR values for different reconstructions of iodine inserts.

FBP	OS-SART	OSTR
4.47	6.99	9.02

Our anecdotal results using real scanner data and an iodine phantom show that OSTR outperformed (in an SDNR sense) the more conventional FBP and SART algorithm, the increased performance is likely due to the fact that OSTR includes an approximate likelihood model as well as edge-preserving regularizer.

Future work will include a study of patient motion artifacts, not studied here since the phantom was static. Preliminary results (Chen et al., 2013) using SART and FBP showed the advantage of RS combined with interleaved acquisition, but additional benefits of a penalized likelihood reconstruction need to be investigated.

Chapter 5 A Closer Look at SPECT

In this chapter, we take a closer look at SPECT imaging and give enough background detail so that work reported in Chapter 6 will be understandable. We have explained SPECT at a rudimentary level in Chapter 1; here we give more detail.

5.1 Overview

In SPECT (Wernick and Aarsvold, 2004), the patient is injected with a radiopharmaceutical tagged with a radionuclide that emits γ-rays in the 50-350 keV range. The γ-ray energy is large enough to pass through the body but low enough to get intercepted by the detector. Its half-life is typically a few hours, long enough for imaging and short enough so it doesn't stay in the patient for days. These requirements mean that there are only 6 radionuclides useful in SPECT (Cherry et al., 2012). The radiopharmaceutical is designed to go to specific organs, vasculature or tumors. The pharmaceutical can express cardiac perfusion, metabolism, cell receptor density or physiologic kinetic behavior. The reconstructed SPECT image is a 3-D map of radionuclide spatial density and can be used in cardiology, oncology, brain function research and other areas.

The raw data is obtained by a gamma camera as explained in Chapter 1. Figure 5.1 shows a typical gamma camera in a clinical setting. The scan time is roughly 20 minutes for most studies. As in DBT the camera swings about the patient, but through 180 or 360 degrees, so the angular sampling for SPECT is not limited. In recent applications in cardiac SPECT, stationary cameras with a limited angular view have found clinical use (ICRU, 1989, Hudson and Larkin, 1994). The camera data are, roughly speaking, line integrals of the radionuclide density, and tomographic reconstruction is used to obtain the SPECT image. Figure 5.2 shows a SPECT image of a brain. Note the low resolution nature of the image. In recent years, both SPECT and PET have been combined with CT scanners so that the functional image of SPECT can be superposed on the high resolution anatomical map provided by CT (Gaemperli et al., 2007, Mariani et al., 2010, Beyer et al., 2000). For SPECT, CT can help localize lesions in the anatomy

and also be used to compute attenuation factors as discussed below. Figure 5.1 shows a SPECT/CT scanner.

Figure 5.1 SPECT-CT scanner. The patient sits underneath the two SPECT camera units. The CT unit is the large circular structure. After the SPECT scan, the patient is imaged by the CT. Photo is in curtsey of Siemens Healthcare.

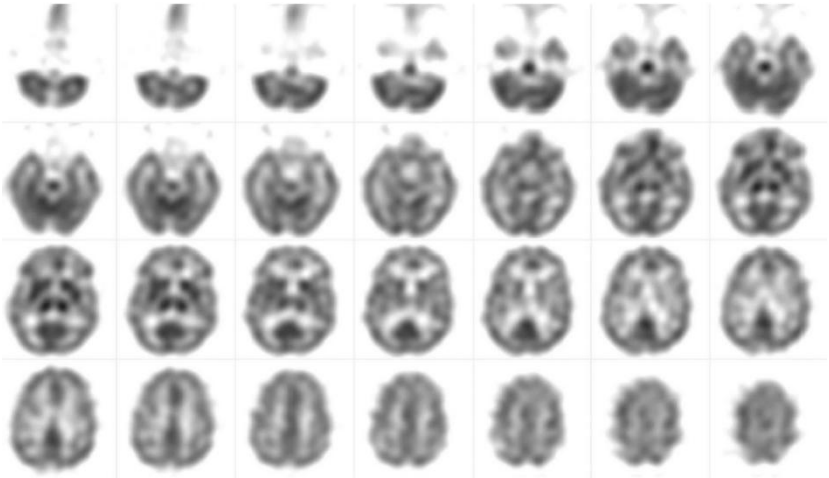

Figure 5.2 Transaxial SPECT images showing perfusion in the brain of a normal adult following injection of 890 MBq of 99mTc-HMPAO. (Courtesy Dr. Steve Meikle, University of Sydney, Australia.)

For many studies, a local increase or decrease in radionuclide density is medically significant. In Chapter 6, our clinical setting is imaging of neuroendocrine tumors (NETs) (Cadiot et al., 1997, Sundin et al., 2007). One problem in NET is to examine the liver for local concentrations of the radiopharmaceutical. For NETs, the radiopharmaceutical is Octreoscan which seeks somatostatin receptors. It is tagged with the radionuclide In-111. These local concentrations can indicate tumors. The problem is thus one of detection and localization: View the liver and search for liver candidates and their locations. This same detection/localization task applies to other areas of SPECT oncology and SPECT cardiology.

SPECT is a true photon-counting modality. The data is the number of photons received in each detector pixel. Count levels for a full study covering, say, the torso could be about 20M, yielding low numbers of photons (0 - 100) per detector pixel. A detector pixel is typically about 3 mm square and a camera face is discretized to 128 x 128 pixels. The number of angular positions of camera is also about 128 in order to satisfy angular sampling criteria for reconstruction.

5.2. Basic imaging chain

The radionuclide emits γ-rays from within the body. As they head toward the camera, they see tissues with very different attenuation coefficients, e.g. lung, soft tissue and bone. The rays are attenuated by Compton and photoelectric interactions as in DBT. This attenuation is a side effect, we are interested in the radionuclide density not the attenuation, but we need to know how the rays were attenuated in order to do a SPECT reconstruction. The γ-rays undergoing photoelectric interactions disappear, but the Compton scattered photons can make it to the detector and get counted as data. In SPECT, attenuation occurs along a ray path starting with the voxel (within the patient) and ending at the detector pixel. This differs from transmission imaging where the attenuation path is from the external source (X-ray tube) to the detector. If a detected photon makes it to the detector unscattered, we call it a *primary* photon.

Due to the collimator action described in Chapter 1, the detector measures approximate line integrals of the radionuclide density. The collimator is a key component in controlling image quality and we talk much more about it below. The γ-rays that make it through the collimator hit a scintillator that generates optical photons. These optical photons are picked up by

photomultiplier (PMT) tubes and the position and energy of the γ-ray is estimated from the PMT outputs. The recorded counts are then used as an input to a reconstruction algorithm.

5.3 The collimator

The collimator is a key element in SPECT. It controls the noise-resolution tradeoff in the data. In the simple model of Chapter 1, the collimator was an array of holes. If we broaden the hole diameter, we get more photons, hence less noise, but do so at the expense of resolution. But this example of a noise-resolution tradeoff is too simple and we consider the collimator in more detail.

We consider a parallel-hole collimator though a variety of other designs of possible (Cherry et al., 2012). Collimators for clinical use are often made of lead. Figure 5.3 gives a rough feel for the thickness and hole size. The collimator is about 50 cm in diameter, enough so its field of view covers the body. Let's take a closer look at a collimator. The hole shape and its lattice arrangement is one factor. (We will be using hexagonal hole shapes in hexagonal arrays). The bore length is the length of the hole and the bore diameter its width. For non-circular collimators, one needs some convention on width and in Figure 5.4 we show a hole diameter for a hexagonal cell. The walls separating the holes are the septa and the septal thickness shown in Figure 5.4 is an important factor.

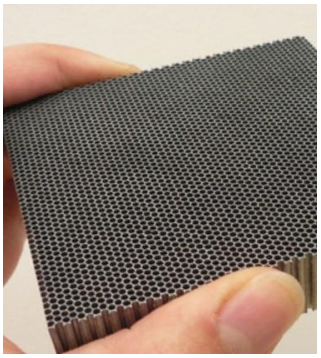

Figure 5.3 A section of collimator. Courtesy of Nuclear Fields.

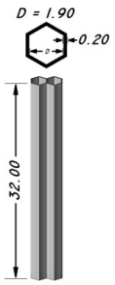

Figure 5.4 Two bores of a parallel-hole collimator illustrating the geometrical parameters. Bores are drawn to scale for a real collimator.

As a first approximation, assume that the collimator has infinite attenuation i.e. blocks any γ-rays that hit the septa. Then from Figure 5.5 we can see that the acceptance angle varies with depth (source to collimator distance). If, for an array of such bores, we plot the pattern of accepted counts on the detector face vs. depth, we get the plot of Figure 5.6. Figure 5.7 shows the origin of this blurring. This depth-dependent resolution can be modeled as a 2-D Gaussian whose standard deviation grows linearly with depth and whose amplitude declines with depth. This depth dependent resolution depends on the hole shape, bore length and bore width, but depends weakly on the septal thickness. (In Chapter 6 we give expressions for the geometric resolution of the collimators). A characteristic resolution distance can be obtained using a point source at some reference distance (typically 15-20 cm) from the collimator face. This resolution distance is the width of the psf due to the point source. (There is a reciprocity of the psf as measured in the detector space or the object space).

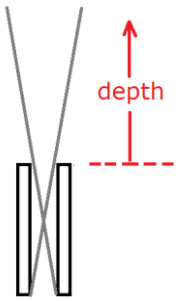

Figure 5.5 Depth dependent resolution for a single bore. From simple ray tracing, the resolution in object space increases linearly with depth.

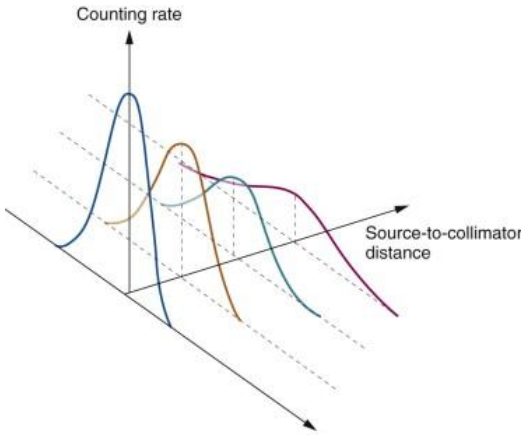

Figure 5.6 The collimator response as a function of depth. Courtesy of (Cherry et al., 2012).

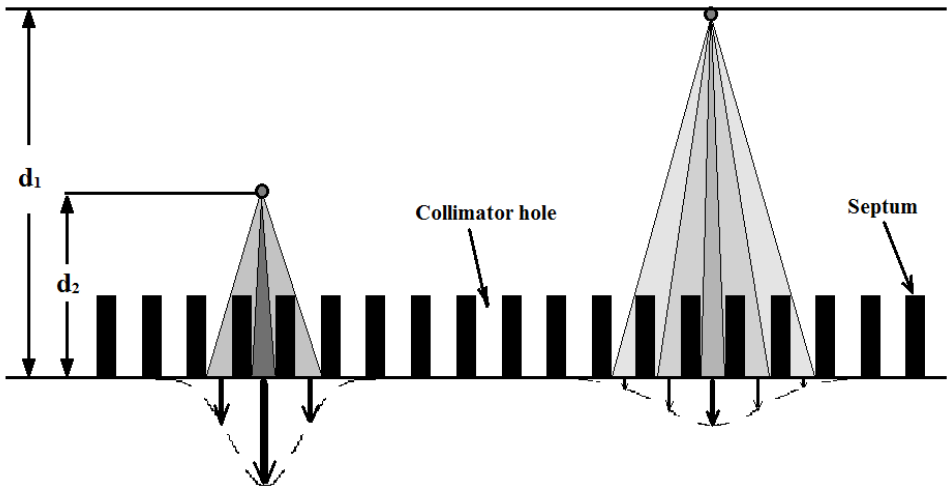

Figure 5.7 Points at different depth yield different point-spread functions on the detector space. The solid angle between septa as seen by the point is proportional to the number of counts received between those septa.

The efficiency of this geometric collimator is due to solid angle effects. Efficiency is defined as the probability that an emission from a point source will get through the collimator. For a hexagonal collimator, (Zhou and Gindi, 2009) efficiency is given by

$$Efficiency = \frac{2}{\sqrt{3}} \frac{\pi D^4}{64 l^2 \left[\frac{\sqrt{\pi}}{\sqrt[4]{12}} D + SPT \right]^2}$$

The efficiency is determined by bore diameter *D*, bore length *l* and septal thickness (*SPT*). Non-circular bore profiles needs a definition of 'bore diameter' and 'septal thickness'. (Gunter, 2004) defined (area of the hole) = $\frac{\pi}{4}D^2$. We use a hexagonal bore shape, then the relation leads to $D = 1.819S$, where S is the length of one of the hexagonal sides. The septal thickness *SPT* is determined by the bore diameter *D* and the hole separation *HOLSEP*, where *HOLSEP* is defined as the distance between the centers of adjacent bores in the lattice. We shall use a hexagonal lattice of hexagonal bores. In this case, we have (Gunter, 2004) $SPT = HOLSEP - \frac{\sqrt{\pi}}{\sqrt[4]{12}}D$.

Interestingly, efficiency is independent of depth for a parallel-hole collimator. Things get more complicated if we consider the fact that the septa must be made thin enough to not block too much radiation by taking up too much area with lead. With thin septa, we must take into account the fact that the γ-rays can pass *through* the septa and get counted as seen in Figure 5.8. This effect is called septal penetration. The reason for this is that when a septum is thin, photoelectric absorption of the γ-ray photons in it cannot stop all the photons. The γ-ray can also Compton scatter in the septa leading to collimator scatter as seen in Figure 5.9. It would appear that photons recorded due to collimator scatter and penetration would be a small percentage of those that pass through the holes without collimator interaction, but surprisingly, collimator effects (scatter and penetration from the septa) can account for a *majority* of recorded counts! This fact plays a key part in our work on collimator design.

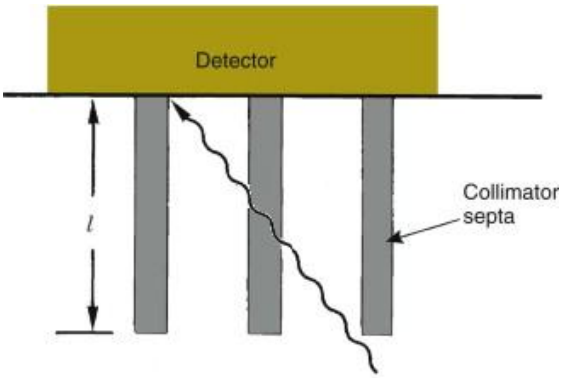

Figure 5.8 Septal penetration.

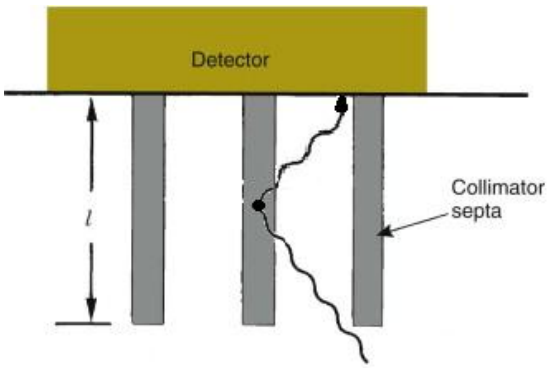

Figure 5.9 Septal scatter.

Collimator effects will cause the psf to widen and exhibit non-isotropy. The spokes on the psf in Figure 5.10 are due to collimator effects for a hexagonal collimator. Efficiency rises as the number of collimator scattered + penetrated photons increase. A key question will be: Do these extra photons carry useful imaging information?

The collimator can also emit characteristic radiation (similar to the X-ray anode) termed "lead K X-rays" at 72 and 75 keV which in some cases can get counted by the detector. In addition, some γ-rays can pass the collimator and then scatter from the detector behind the crystal back into the detector and get counted in an incorrect location. These "backscattered" photons can add additional counts.

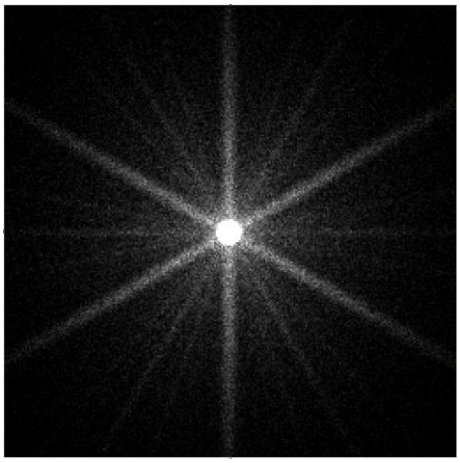

Figure 5.10 The spokes in the collimator psf are due to septal penetration in the hexagonal bore.

5.4. γ-rays detector - the Anger camera

The detector system for SPECT was developed in the 1950's by Hal Anger (Anger, 1958). Once through the collimator, the γ-ray hits a scintillator, almost always a large NaI(Tl) crystal that is about 1 cm thick and 50 cm in diameter. The γ-rays hit fluorescence centers and generate a shower of optical photons mainly in the blue end of the spectrum. These photons spread though a light diffuser to an array of photomultipliers. Figure 5.11 shows a PMT array. The number of optical photons per γ-ray is random. A further amount of uncertainty is introduced when one optical photon hits a PMT. An optical photon generates a random number of photoelectrons. The number of photoelectrons are amplified (the amplification has surprisingly low noise) through a dynode chain to deliver a current pulse at the output of the PMT.

Now the PMT outputs are doubly random due to the random number of optical photons per γ-ray and random number of photoelectrons per optical photons. PMTs closer to the scintillation event get more photons, and generate more current output. The relative outputs of the few PMTs can then be used to estimate the position of the scintillation event with surprising precision. The PMT outputs are noisy, so position estimation is a statistical estimation problem. This statistical estimate of position is beyond the scope of our work though.

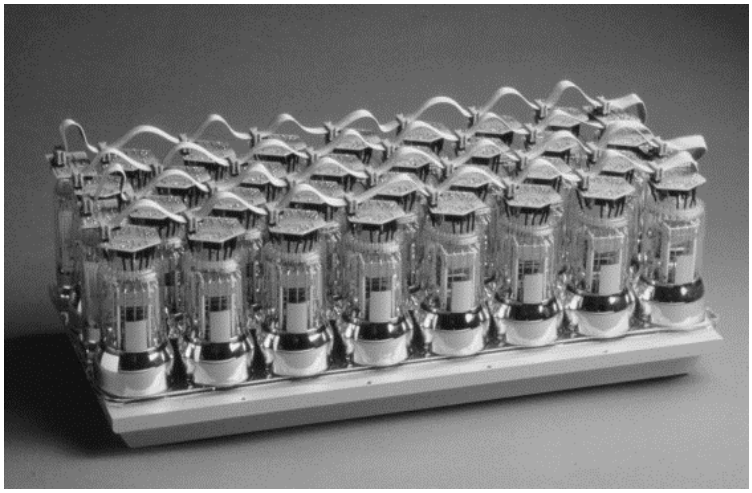

Figure 5.11 PMT array used in a gamma camera. Courtesy of (Cherry et al., 2012).

The estimated position of the scintillation event is discretized and, if the event is in an acceptable energy range (to be discussed below), then the number of counts in one detector "bin" (memory location corresponding to a discretized location) is incremented by one. This is how the camera image is formed.

Note that since position estimation is noisy, the detector has an intrinsic spatial resolution due to counting statistics. That is, if we took an infinitely thin "pencil beam" of radiation directed at the crystal with no collimator in the system, the decoding circuitry would give a variety of positions for each γ-ray in the beam, even though the camera position where the beam hits is fixed. The accumulated image would be approximately a spatial blob, typically of FWHM (full width half max) of about 3 mm. This distance is termed the intrinsic resolution of the camera. Therefore the collimator and the camera contribute spatial resolution loss. For a point source in air at distance d from the camera face, the collimator resolution and intrinsic resolution combine in quadrature $\sigma = \sqrt{(\alpha d)^2 + \sigma_0^2}$ where α = a constant, σ_0 is the intrinsic resolution and d the depth.

The "Anger camera" described above is odd in that there are no physical detector pixels. Instead, the position and energy of an event is obtained from the outputs of a coarse array of PMT tubes, and the position estimate is quantized to a particular detection "bin" which is equivalent to a pixel.

In recent years, after many years of research, solid state detectors have finally found use in SPECT and have already found commercial application in cardiac SPECT systems (ICRU, 1989, Hudson and Larkin, 1994). These use CZT (Cadmium Zinc Telluride), are pixilated and have better spatial and energy resolution than the Anger camera. The Anger camera is still the workhorse in SPECT and I use Anger camera simulation in my work.

5.5 The energy spectrum of the data

The γ-ray photons generated by the radionuclide have a well defined energy or energies. For example, In-111, the one we model, radiates photons at 171 keV and 245 keV at relative prevalence of 0.905 and 0.94. But via scatter in the body and collimator, the photon energy can change by the time it hits the scintillator. If possible, one would like to reject scattered photons

by filtering those whose energy has dropped below the 171 and 245 keV lines. However, due to finite energy resolution of a camera, this rejection will not be perfect. The energy resolution of the detector determines the rejection efficiency with which this can be accomplished.

The energy resolution largely depends on the statistical fluctuations in the number of light photons collected from a scintillation event. The energy resolution varies approximately in proportion to $1/\sqrt{E}$. Thus, it is very hard to separate those small angle scattered photons (with relative high energies) from the nonscatter photons. Thus a energy acceptance window (*photopeak window*), ±10% of the peak energy is usually chosen, i.e.we only count photons within the photopeak windows.

Figure 5.12 shows a typical energy spectrum for an In-111 clinical study. Ideally the spectrum would be two spikes at 171 and 245 keV. However, the detector energy resolution broadens those peaks into the two lumps seen. The background contributions outside the photopeak windows is mainly due to patient scatter with a smaller contribution due to collimator scatter, lead X-rays and backscatter.

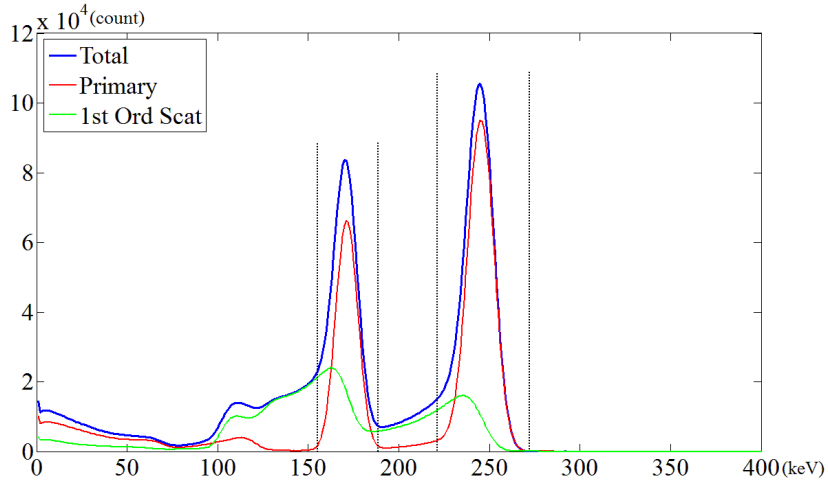

Figure 5.12 Energy spectrum for a typical In-111 SPECT study showing the contributions of all photons and the contributions by single-scattered photons. Vertical dotted lines delineate the ±10% photopeak energy windows.

As can be seen in Figure 5.12, some scatter counts appear within the photopeak window and can't be easily rejected. Some of these are "downscatter" due to patient Compton scatter of the

245 keV γ-ray into the 171 keV photopeak. Because scattered photons are redirected, they lose imaging information (same as in DBT) and lower image contrast. Considerable effort has gone into patient scatter correction for SPECT (Hutton et al., 2011), but we do not discuss that topic here. Our own lab has investigated patient scatter in previous papers (Kulkarni et al., 2009).

5.6 The imaging equation for SPECT

The radionuclide "amount" in a voxel is characterized by the number of counts emitted per unit time. But count emission follows Poisson statistics, so we use the *mean* count rate f_n to measure the "amount" of radioactivity in voxel *n*. Thus f_n is the No. photons /4π solid angle/ voxel/ sec. The collection of f_n is 3D lexicographically ordered as vector $\mathbf{f}$. In SPECT, we seek, using reconstruction, an estimate $\hat{\mathbf{f}}$ of $\mathbf{f}$. The integer number of detected counts in detector bin *m* is denoted by g_m and vector $\mathbf{g}$ accounts for all bins at all angles. (So here, $\mathbf{g}$ plays a role analogous to its role in DBT).

A photon leaving voxel *n* has a probability H_{mn} of arriving at bin *m*. The probability is small due to absorption and scatter in the patient, collimator scatter and absorption, and due to the very limited solid angle of acceptance by the collimator. As a probability, a typical matrix element $H_{mn} \sim 10^{-4}$. Therefore, SPECT is a count-starved modality. (Actually, for modeling, H_{mn} needn't be the actual probability, it only needs to be proportional to the probability to not affect the reconstruction.)

In Chapter 6, we will need the likelihood for SPECT detector counts. We make use of derivations from previous work in our lab (Wang, 1997) to derive this. A voxel emits a Poisson number of photons in an exposure time *T* into 4π steradians, and the mean number of photons is f_n. If one has a Poisson random variable of mean α and each event (photon) is then selected by a Bernoulli process with probability p, the random variable associated with the compound Bernoulli-Poisson process is still Poisson with mean $p\alpha$.

We can think of H_{mn} as the Bernoulli probability that a photon from voxel *n* reaches detector bin *m*. Let G_m be the random variable associated with the number of counts in bin *m*. Then by the above, $G_m \sim Poisson\left(H_{mn} f_n\right)$.

Now let's say we have two voxels *n* and *n*' viewed by one detector bin at *m*. The bin receives a Poisson number of counts, with mean $H_{mn} f_n$ from voxel *n* and $H_{mn'} f_{n'}$ from voxel *n*'. It can be shown that the sum of two Poisson random variables is still Poisson but the means add. We can extend this argument to $n = 1,...,N$ voxels seen by a single detector *m*. The counts in that detector are then distributed as $G_m \sim Poisson\left(\sum_{n=1}^{N} H_{mn} f_n\right)$.

Now let's say one has a single voxel *n* viewed by two detectors *m* and *m*'. The photon counts received by each detector are characterized by a Poisson random variable. They will each receive a Poisson number of counts with means $H_{mn} f_n$ and $H_{m'n} f_n$, respectively. It is not obvious, but it can be shown that these two random variables are *independent* (Wang, 1997).

We can extend the above argument to $n = 1,...,N$ voxels viewed by $m = 1,...,M$ detectors. The counts in each bin are Poisson and independent. Let G_m be the random variable for bin *m* and g_m the observed number of counts. Let $\mathbf{G}$ be a random vector $\left(G_1,...,G_m\right)$. Then putting the results together we get

$$\begin{aligned} \Pr\left(\mathbf{G} = \mathbf{g} \mid \mathbf{f}\right) &= \prod_{m=1}^{M} \Pr\left(G_m = g_m \mid \mathbf{f}\right) \\ &= \prod_{m=1}^{M} \frac{\left(\sum_{n=1}^{N} H_{mn} f_n\right)^{f_n} \exp\left(-\sum_{n=1}^{N} H_{mn} f_n\right)}{g_m!} \end{aligned} \tag{5.1}$$

If we revert to a more compact notation using only $\mathbf{g}$ and $\mathbf{f}$ and consider that $\mathbf{g}$ is known, then the likelihood of $\mathbf{f}$ is

$$\Pr\left(\mathbf{g} \mid \mathbf{f}\right) = \prod_{m=1}^{M} \frac{\left[\mathbf{Hf}\right]_m^{g_m} \exp\left(-\left[\mathbf{Hf}\right]_m\right)}{g_m!} \tag{5.2}$$

where $\mathbf{Hf} \equiv \sum_{n=1}^{N} H_{mn} f_n$. In Chapter 7 we shall use the notational convention of eq.(5.2).

Given the likelihood for SPECT, we can use the same arguments as for the statistical reconstruction for transmission in Chapter 4 to write SPECT reconstruction as maximization of a penalized likelihood

$$\hat{\mathbf{f}} = \arg\max_{\mathbf{f}} \Phi(\mathbf{f}) = \sum_m \left[g_m \log\left[\mathbf{Hf}\right]_m - \left[\mathbf{Hf}\right]_m \right] - \beta \cdot \sum_n \sum_{k \in \mathrm{N}(n)} \frac{1}{2} \omega_{n,k} \Psi\left(f_n - f_k\right) \tag{5.3}$$

where the log likelihood term takes a form consistent with eq.(5.2). In this thesis, we shall not be concerned with reconstruction aspect of SPECT but I include eq.(5.3) just for complete description. Our lab has had extensive experience in researching SPECT reconstruction.

The SPECT imaging equation takes the simple linear form

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n} \tag{5.4}$$

with $\mathbf{n}$ Poisson since $\bar{\mathbf{g}} = \mathbf{Hf}$.

Given $\mathbf{f}$ in the form of a digital phantom, we can calculate $\mathbf{g}$ using a Monte-Carlo (MC) package. The MC package generates photons at voxel *n* with probability governed by f_n and tracks the photons by Monte-Carlo simulation of the physics of random directional emission, photoelectric and Compton effects in the patient, and interactions in the collimator and scintillators. (In our MC version described in Chapter 7 shortcuts are taken to model the complex PMT) By tracing enough source events, we implicitly carry out eq.(5.4). Thus the matrix product $\mathbf{Hf}$ is not explicitly computed but carried out by a complex Monte Carlo package that tracks photons. I note that in reconstruction algorithms, the $\mathbf{Hf}$ product is usually carried out by an on-

the-fly approximate matrix vector product that is not as accurate as MC but is more practical computationally.

We have given enough background in SPECT so we can discuss work to date in Chapter 7. We have not talked about tasks and observers for our SPECT project since that is contained in Chapter 7.

Chapter 6 Task Based Approach to SPECT Optimization

6.1. Introduction

In previous chapters, crude empirical measures such as SDNR were used to quantify the "goodness" of images. A more rational approach (Barrett and Myers, 2003) is to mathematically define the medical task, an observer to carry out the task, an observer response and finally a means to compute a scalar task performance figure for purposes of comparing or optimizing imaging systems. In Chapter 7 we will use this approach to select an optimal collimator for SPECT NET imaging. In this chapter, we give preliminary material on task-based image quality assessment to support the research reported in Chapter 7.

Assume for a moment that our medical task was to decide whether a signal (e.g. tumor) was present or absent in a medical image. This is a simple 2-class detection problem. We give an observer, which could in theory be a human observer, a series of images in which the signal is known present or known absent. The observer, who has no knowledge of the presence/absence of the signal, has to decide presence or absence. If the signal is decided present when it is in fact present and absent when it is, in fact, absent, then the performance is quite good.

Now let's say we repeat this performance analysis for a series of different collimators, or perhaps reconstruction parameters. By choosing the collimator or reconstruction with the best performance figure on the detection task, we have optimized the imaging system in a rational task-dependent way.

In practice, the labor involved in using human observers (e.g. a panel of radiologists) is too time consuming for practical system optimization. We can replace the human by a *mathematical* observer that reports a scalar observer response. By comparing this response to a threshold, a binary decision signal "present" or "absent" is made. If a human is to be replaced, we seek a human emulating model (i.e. mathematical) observer. For the binary decision task there has been much research on the design of such observers, but we will not take advantage of this work.

We have replaced the human observer by a model observer, but for our task we also need to replace the simple detection task, in which a signal at a known location is deemed present or

absent, by a more complex task. Our medical problem involves inspection of the liver to find the possible presence of an NET. This is a joint detection/estimation problem: the observer must look at the liver, decide "yes" or "no" on whether an NET is present, and if "yes", estimate *where* the lesion is present. Again, one might use a human-emulating model observer to examine the reconstructed image for the purposes of optimizing some aspect of the imaging system. Indeed, such model observers have been proposed, but in our work we take a different approach.

Our goal is to optimize the collimator. The collimator delivers projection date $\mathbf{g}$, which is subsequently reconstructed to deliver a human-viewable reconstruction $\hat{\mathbf{f}}$ as described previously. Indeed, the information needed to support the task is implicitly contained in $\mathbf{g}$ and reconstruction may be viewed as a means of "reformatting" $\mathbf{g}$ into a form $\hat{\mathbf{f}}$ more suitable for human viewing. If this is the case, it is simpler to apply a model observer to $\mathbf{g}$ directly rather than a human-emulating model observer to $\hat{\mathbf{f}}$.

More generally, in any tomographic system, "raw" projection data is first collected by the imaging hardware and this data is them reconstructed by a computer algorithm for purposes of human viewing. In the image-quality community, for purposes of optimizing hardware in tomographic systems, the consensus is that it is best to apply a model observer to the raw data rather than a human-emulating model observer to the reconstruction. One reason for this is that the math will be easier if we operate the $\mathbf{g}$ space.

If a non-human-emulating observer is to be applied to the raw data, what observer should we use? The consensus is to use the *ideal observer*, which is the one that delivers the best possible performance metric of all observers (Barrett and Myers, 2003). That is, armed with the pdfs characterizing $\mathbf{g}$, the ideal observer will take maximum advantage to optimize performance.

The questions that remain then are: (i) What is the ideal observer for our task? (ii) How is performance computed? As a preliminary clue to the answer to these questions, we consider the analysis of the 2-class detection problem as formed in standard texts such as (Van Trees et al., 2013). The ideal observer turns out to be the log-likelihood ratio, the performance involves computation of the ROC (receiver operating characteristic) curve, and the final scalar performance is AROC, the area under the ROC curve. We discuss these and generalize them to our more complex task below.

One last non-intuitive point is this: Even for a 2 class signal present/absent task, the presence of a signal in the raw data does not show up as a localized blob; instead, the signal is distributed widely in $\mathbf{g}$. The ideal observer is still able to take the data and optimally infer the presence and location of the signal in the $\mathbf{f}$ space by inspecting the $\mathbf{g}$ space. This is true despite the fact that to a human, there is no easily seen compact "blob" in the $\mathbf{g}$ space. We illustrate this in a bit more detail below.

6.2 Ideal Observer

In previous work in our lab, we formulated the general ideal observer (IO) for the detection-localization task (Khurd and Gindi, 2005a). Here, we review the IO and refer the reader to (Khurd and Gindi, 2005a) for further details on its formulation. We define the IO in the context of SPECT here, but in (Khurd and Gindi, 2005a) its formulation is general, and applies to any data processing problem.

Let the source activity be described by vector $\mathbf{f}$ with components $f_n, n = 1,..., N$ with f_n being the activity (mean counts/sec/s.r./voxel) in voxel n and N being the total number of voxels. In a real imaging system $\mathbf{f}$ is continuous, but here we use a voxelized version appropriate for our experiments. Let vector $\mathbf{g}$ $\{g_m, m = 1,..., M\}$ denote the camera bin counts in bin m, with M the total bin number in all camera faces at all angles. Then the imaging equation is

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n} \tag{6.1}$$

where $\mathbf{H}$ is the $M \times N$ system matrix with element $\mathbf{H}_{mn}$ proportional to the probability that a photon from voxel n is registered at bin m, and $\mathbf{n}$ is the Poisson noise with n_m an instance of the noise in bin m. We use the Monte Carlo package SIMIND to compute $\mathbf{Hf}$ and then add Poisson noise $\mathbf{n}$.

Our task is to detect whether a signal is present somewhere in a search region Ω. Our search region, located in the liver, will be described in more detail below. Let $\mathbf{s}_j$ be a signal of known form at location $j \in \Omega$. Let $p(\mathbf{s}_j)$ be the prior probability that the signal will be found at j. We will use a uniform $p(\mathbf{s}_j)$ without loss of generality. The notion of localization is meaningless

without specifying a tolerance distance about the true location within which localization is deemed correct. That is, a reader who localizes a lesion at a slight distance from the true location is said to have correctly localized it. This distance (tolerance) depends on the clinical context. Our test results are evaluated at several tolerance distances. Let $l = 1,...,L$ also index signal location. We define the hypothesis H_l, where $l = 1,...,L$, to mean that the signal is located within the circular tolerance region $T(l)$ centered at l. The hypothesis H_0 means signal absent. Thus, we have an $L+1$ hypotheses decision problem. We shall also define a generalized likelihood ratio $\mathrm{LR}(\mathbf{g},\mathbf{s}_j)$, that is indexed to the signal at j, as:

$$\mathrm{LR}(\mathbf{g},\mathbf{s}_j) = \frac{p(\mathbf{g}\,|\,H_j)}{p(\mathbf{g}\,|\,H_0)} \tag{6.2}$$

With these definitions, we can write our optimal decision strategy for detection and localization:

$$\begin{aligned} t(\mathbf{g}) &= \max_{l\in\{1,...,L\}} \sum_{j\in T(l)} p(\mathbf{s}_j)\mathrm{LR}(\mathbf{g},\mathbf{s}_j) \\ l(\mathbf{g}) &= \arg\max_{l\in\{1,...,L\}} \sum_{j\in T(l)} p(\mathbf{s}_j)\mathrm{LR}(\mathbf{g},\mathbf{s}_j) \\ &\textit{Decide } H_{l(\mathbf{g})} \textit{ if } t(\mathbf{g}) > \tau, \textit{ else decide } H_0 \end{aligned} \tag{6.3}$$

Our decision strategy will detect the presence or absence of the signal in the data $\mathbf{g}$ by comparing a scalar observer response $t(\mathbf{g})$ to a threshold τ and, if a signal is detected, it will report a location $l(\mathbf{g})$ where it deems the signal to be present. We shall use 'tolerance' to simply indicate the diameter of $T(l)$, and also note that in our experiments, the tolerance does not change with position. Thus the IO in (6.3) examines $\mathbf{g}$ and chooses the maximum of the generalized likelihood ratios (after convolution with the tolerance template) as the observer response $t(\mathbf{g})$. The IO claims signal present if $t(\mathbf{g})$ exceeds a decision threshold. The location where the max occurs is also the estimated signal location $l(\mathbf{g})$. If the signal is deemed present, then $l(\mathbf{g})$ is the estimated location.

From the likelihood equation for SPECT, eq.(5.2), one can easily express the likelihood ratio in (6.2) as:

$$\mathrm{LR}(\mathbf{g},\mathbf{s}_j) = \prod_m \exp(-[\mathbf{Hs}_j]_m)\left(1+\frac{[\mathbf{Hs}_j]_m}{[\mathbf{Hb}]_m}\right)^{g_m} \tag{6.4}$$

Equations (6.1) to (6.4) define the IO. The reconstruction $\hat{\mathbf{f}}$ of $\mathbf{g}$ does not enter into the calculation. Also note that one is tempted to take the log in (6.4) to simplify the right side in (6.4), but this is illegal mathematically due to the "max-sum" formulation in (6.3).

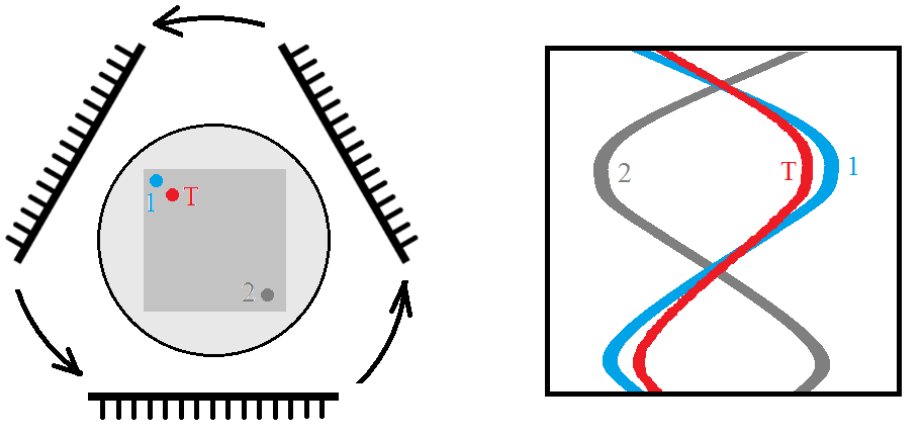

Figure 6.1 Intuitive explanation of ideal observer action. The simple 2D SPECT experiment shows the object at left, with a true signal T and two hypothesized signals 1 and 2 within a search region denoted by the grey rectangle. The corresponding sinogram is shown at the right. The IO finds that sinogram that best matches the true-signal sinogram and considers the corresponding location in the object space as the best guess.

The action of the IO is somewhat abstract, and Figure 6.1 uses a simple cartoon of 2D SPECT to clarify. We make use of the concept of a "sinogram". For a 2-D object, each projection is 1-D. We can stack the projections with angle as the ordinate and distance as the abscissa in Figure 6.1. If the object is a single point, the sinogram is a sinewave whose amplitude and phase encodes position and whose intensity encodes intensity in the object space. In this sense, reconstruction is a process of converting a sinogram to an object estimate. In our case, if the problem involved a 2-D object, then the IO would inspect the sinogram $\mathbf{g}$, not the reconstruction.

For a 3-D object, we have many 2-D camera faces as will be graphically illustrated in Chapter 7. These faces comprise $\mathbf{g}$ and will be viewed by the IO.

Basically the IO examines the sinogram to make inferences about the object. In Figure 6.1, the true signal is located at "T" and two hypothesized signal locations are shown at "1" and "2". The sinograms for all three are shown. Essentially, the IO takes the sinogram of hypothesized locations (1 and 2 in this case) and computes a measure of closeness to the true sinogram for each. The location of the signal where hypothesized sinogram that is closest ("1" in this case) to the true sinogram is selected as the best candidate and the signal is deemed present if the closeness measure exceeds a threshold. For 3D SPECT, the "sinogram" is the collection of the 60 camera face images. While this explanation is meant to provide simplified intuitive insight, equations (6.3) and (6.4) provide an exact mathematical definition of the IO.

6.3 Computing ALROC, the Performance Figure of Merit

The IO is optimal in several senses (Khurd and Gindi, 2005a) but in particular it maximizes the ALROC, the area under the LROC (Location Receiver Operating Characteristic) curve (Swensson, 1996). The ALROC is our performance figure of merit.

To explain ALROC and the method for its computation, it is instructive to first revisit the more familiar ROC (receiver operating characteristic) curve (Metz, 1978, Van Trees et al., 2013). The ROC curve is appropriate for a detection-only task. The ROC curve plots the true positive (TP) rate vs. the false positive (FP) rate of detection by an observer that makes a decision based on an observer response compared to a threshold. Here the TP rate is the threshold-dependent probability of stating that a signal is present when it actually is present, and the FP rate is the threshold dependent probability of stating that a signal is present when it is absent. The curve is swept as the threshold varies. The area under this curve, AROC, is deemed a scalar performance figure of merit. In (Metz, 1978), it is explained that the ROC curve can be generated by first forming suitably normalized histograms of observer responses for signals truly present and signals truly absent, then integrating each curve at values above the threshold. As the threshold is swept, the values obtained by integration sweep out the ROC curve.

The LROC curve is similar (Zhou et al., 2008), but plots the correct localization (CL) rate vs. the FP rate. A signal is correctly localized if it is truly present and deemed present by the

observer ($t(\mathbf{g}) > \tau$ in Eq.(6.3)), and furthermore the estimated location $l(\mathbf{g})$ is within the tolerance radius about the true signal. The LROC curve can also be computed by a histogramming procedure. For a given collimator corresponding to a given $\mathbf{H}$ and many signal-present data sets and many signal-absent datasets, compute the two signal-present and -absent histograms of observer responses. For the signal present histogram, exclude responses corresponding to mislocalized signals, i.e. ones where $l(\mathbf{g})$ was outside the tolerance region. Normalize the histograms and integrate each as a threshold is lowered to sweep out the LROC curve. Then integrate the curve to obtain ALROC.

While the above description is qualitative, further technical detail on computing ALROC can be found in our previous work (Zhou et al., 2008) (Lu, 2011). Note that this method of computing ROC and LROC curves is appropriate for a numerical observer such as the IO. For human observers who report a "confidence level" instead of an observer response, elaborate fitting procedures are needed to estimate the ROC or LROC curve. Below, we give specifics on our ALROC computation.

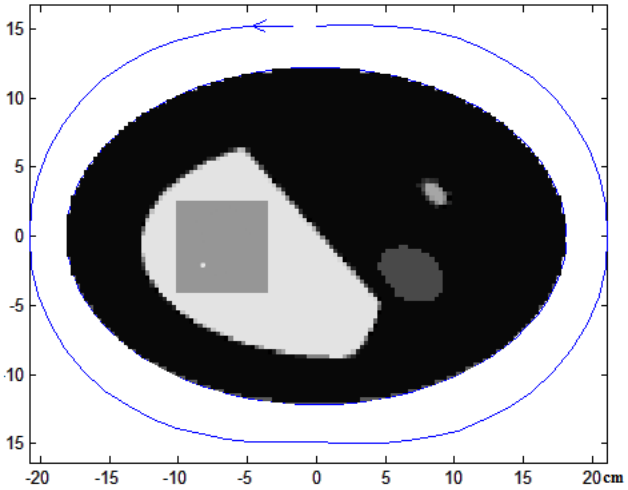

Figure 6.2 One slice of the 128x128 MCAT phantom. The dark gray square in the liver (large organ) shows the 21x21 search region. A one-pixel tumor at 28:1 contrast is shown in the search region. The arrow line shows the contour-following collimator orbit.

The liver search region was confined to a 2D slice that comprised a rectangular region of 21 by 21 pixel locations where a tumor (i.e. a "signal") could be centered. The search region Ω is

depicted in Figure 6.2.We used a signal of one pixel with a contrast chosen so that our observer performance values (ALROC) were not pegged too low nor too high for comparisons between collimators. The signal diameter and contrast is not realistic, but in task performance studies it is common to choose signal (tumor) contrasts appropriate for the task, in this case a small limiting size and appropriate contrast. True signals can occur anywhere in Ω, but it is not necessary (Zhou et al., 2008) to sample all locations in Ω. That is, Ω contains all hypothesized locations that we test, but we only need to generate signal-present camera face data $\mathbf{g}$ for a subset of signal in Ω. Hence, it is sufficient to place true signals in a signal grid S that is a subset of Ω. For S we used a 5 by 5 signal grid centered in Ω with each location spaced 2 pixels from, its nearest neighbor**.** The rectangular search grid shown in Figure 6.2 does not follow the liver contours but its shape is convenient for reducing computational complexity and does not alter the ALROC results. Our final results are thus reported as ALROC at a given tolerance vs. collimator choice.

The specific steps for computing ALROC follow (Lu, 2011). Below we list the mechanical steps in obtaining ALROC but refer the reader to (Zhou et al., 2008) and (Khurd and Gindi, 2005b) to obtain a more detailed insight. We first generate N_{samp} signal-present camera face images $\mathbf{g}_+^k, k = 1,..., N_{samp}^+$. To do this we need to generate N_{samp}^+ signal-present objects $\mathbf{f}_+^k = \mathbf{b} + \mathbf{s}_i^k$, $i \in S$, where $\mathbf{b}$ is the fixed background object and $\mathbf{s}_i^k$ is the kth true signal located at pixel i. We use SIMIND to compute $\mathbf{g}_+^k = \mathbf{H}\mathbf{f}_+^k + \mathbf{n}^k$ where $\mathbf{H}$ is chosen to model a particular collimator. Here $\mathbf{n}^k$ is the kth realization of Poisson noise. The signals $\mathbf{s}_i^k$ are sampled according to a uniform distribution $p(\mathbf{s}_i)$. We also generate N_{samp}^- signal-absent camera face images $\mathbf{g}_-^p$, $p = 1,..., N_{samp}^-$ by using SIMIND to compute $\mathbf{g}_-^p = \mathbf{H}\mathbf{f}_-^p + \mathbf{n}^p$ with $\mathbf{f}_-^p = \mathbf{b}$, i.e. no signal present.

After accumulating N_{samp}^+ $\mathbf{g}_+$ and N_{samp}^- $\mathbf{g}_-$, we apply the ideal observer (6.3) and (6.4) to obtain observer responses $t(\mathbf{g}_+^k)$ and $t(\mathbf{g}_-^p)$. The observer responses are accumulated in histograms $T(\mathbf{g}_+)$ and $T(\mathbf{g}_-)$. The signal-present histogram $T(\mathbf{g}_+)$ is converted to a correct-localization histogram $CL(\mathbf{g}_+)$ by expunging those observer responses $t(\mathbf{g}_+)$ whose corresponding location estimate $l(\mathbf{g}_+)$ is outside the tolerance radius. After filtering mislocalizations, the area of $T(\mathbf{g}_+)$ drops from N_{samp}^+ to N_{samp}^{CL}. The histograms $T(\mathbf{g}_-)$ and

$CL(\mathbf{g}_+)$ are normalized so that their areas are unity and $N_{samp}^{CL} / N_{samp}^{+}$, respectively. After the normalization, we integrate $T(\mathbf{g}_-)$ and $CL(\mathbf{g}_+)$ as a threshold is lowered. The integrated values sweep out the LROC curve, and numerical integration of the LROC curve yields ALROC.

We used $N_{samp}^{+} = 2000$ and $N_{samp}^{-} = 1000$, sample numbers high enough to ensure a 2% uncertainty in ALROC (Zhou and Gindi, 2009). We note that $N_{samp}^{+} > N_{samp}^{-}$ since many of the signal-present $\mathbf{g}_+^k$ are excluded due to mislocalization.

We have given a motivation for task-based imaging system optimization, pointed out that the observer for a tomographic system should be the IO viewing the **g** space, given the IO for our detection localization task, and given some detail on computing the task performance figure of merit, ALROC.

Chapter 7 Collimator Performance Evaluation for In-111 SPECT Using a Detection/Localization Task

This chapter presents any detailed work on collimator optimization. The contents of this chapter appeared as (Lu et al., 2014c). My co-authors on this publication were Lin Chen (2nd author) and Gene Gindi (third author). The actual work was done almost completely by me with guidance from my advisor Dr. Gindi.

7.1 Introduction

For imaging neuroendocrine tumors (NETs), single-photon imaging using In^{111}-pentetreotide (Octreoscan) is the method of choice (Cadiot et al., 1997) (Krenning et al., 1994) (Sundin et al., 2007) (Mäahler et al., 2012). Since the area over which NETs can occur is extensive, a large field of view afforded by parallel-hole collimators is needed. A common protocol is to first perform planar imaging over a wide area (two camera frames), then follow this with a SPECT scan. The SPECT allows better visualization of small tumors suspected in the planar scan.

In our clinic at Stony Brook University Medical Center, we image about 50 patients per year. Planar imaging is done post injection followed by a SPECT scan at 24 hours post injection over two frames, one over the chest and one over the abdominal region. Our imaging protocol leads to about 20 M total counts per frame for the SPECT images. We use a Siemens E-CAM camera with an SEME (Siemens ECAM Medium Energy) hexagonal parallel-hole collimator to achieve this count level. The choice of the SEME collimator is motivated in part by the fact that In^{111} radiates mainly at 171 keV and 245 keV at a prevalence of 0.905 and 0.940, respectively. The SEME collimator septa are thick enough to block septal penetration from the higher energy emissions. However, the high-energy septal-penetrating photons may carry useful imaging

information; by eliminating them using thick septa, we may be lowering the useful count level, thus increasing the noise in the reconstruction.

As pointed out in (Mähler et al., 2012), NET imaging involves a relatively short acquisition time, a low amount of injected activity and a post injection delay that is a significant fraction of the radionuclide decay of the In^{111} over the 67.3 hour half-life for In^{111}. These factors result in SPECT data with low counts and a high noise level. (Mähler et al., 2012) modeled a clinically used relatively low efficiency MEGP (Medium Energy General Purpose) collimator that exacerbated the noisy data problem. They proposed to replace the MEGP by an ELEGP (Extended Low Energy General Purpose) collimator that allowed an increased level of septal penetration. Their simulation used a cylindrical physical phantom designed to mimic the activity in liver containing liver metastases. The detection of small NET liver metastases is important clinically and, in particular, in staging neuroendocrine cancer. In (Mähler et al., 2012) and in the present work the focus is on tumor detection rather than quantitation.

In (Mähler et al., 2012) contrast and CNR (contrast-to-noise ratio) were used as figures of merit to assess the visibility of simulated NETs in the liver. These empirical figures of merit were applied to reconstructions derived from iterative algorithms. These figures of merit were the basis for their comparison of both collimators and compensation methods in the reconstruction. With model-based scatter compensation replacing conventional scatter compensation and with the ELEGP replacing the MEGP, they showed that the visualization of NETs was improved.

We are also interested in collimator design for optimization of visibility of NETs in the liver, taking into account resolution, efficiency and penetration of the collimators. However, instead of using an empirical figure of merit like contrast or CNR, we use a much more clinically relevant measure of performance based on a clinically relevant task. One of the radiologist's tasks is to search the liver for NETs and detect the presence and location of NET candidates. We can carry out this detection and localization task and measure its performance by using a novel numerical observer (Khurd and Gindi, 2005a) that we developed. One of the central contributions of this paper is the application of this clinically relevant detection and localization task to the design of SPECT collimators.

Our numerical observer is a computer program that is applied to the camera (projection) data. The observer delivers a pair of scalar observer responses. One response is the "confidence", an

estimate of the probability that a signal (NET) is present, and the other response is an estimate of the location of the signal. That is, an observer response pair is obtained from the projection data, which comprises the camera face images from all angles. By simulating many sets of noisy projection data we obtain a set of response pairs. This set can be used to compute a scalar figure of merit of task performance, the ALROC, (area under the localization ROC curve). (Khurd and Gindi, 2005a)

Assessing task performance on the raw projection data instead of on the reconstructed data may seem counterintuitive. A more common approach (Gifford et al., 2005) (Wollenweber et al., 1999) is to apply a numerical observer to the reconstruction and to design such a numerical observer to emulate the performance of a human observer. However, we instead follow an approach in which assessment is performed on the raw projection data. For system optimization, i.e. optimization of imaging hardware such as a collimator, this approach is advocated strongly in (Barrett and Myers, 2003), and used for SPECT by (Lee et al., 2013) and (Ghaly et al., 2013).

Of all the types of numerical observers that can be applied to the projection data, we use the "ideal" observer as advocated in (Barrett and Myers, 2003). The ideal observer is the one that yields the best possible figure of merit for the task, in this case ALROC. The actual mathematical form of the ideal observer depends on the task. For our detection/localization task, we formulated the general form of the ideal observer in (Khurd and Gindi, 2005a).

We specialized this general form of the ideal observer to the case of SPECT imaging in (Zhou and Gindi, 2009). In (Zhou and Gindi, 2009) we performed a collimator optimization study in a toy-like setting, simulating 2-D SPECT and using a simple geometrical phantom that did not model a clinical problem. We performed image simulation with an analytical projector that ignored the physical effects of patient scatter, septal penetration and septal scatter. Furthermore, the simulations were based on the monoenergetic Tc^{99m} radionuclide (140 keV), and so complex collimator tradeoffs due to multispectral radionuclides such as In^{111} were not considered. In the current study, we use Monte Carlo methods to simulate projection data from a 3D anthropomorphic MCAT Phantom modified for the NET problem. This approach allows us to include a multispectral (In^{111}) radionuclide, patient scatter and septal penetration and septal scatter effects. As we will show in Sec. 7.3, these collimator effects turn out to play an important role in collimator performance.

In Sec. 7.2.1, we describe the details of the modified MCAT phantom used for our NET imaging simulations. We describe the search region in the liver where NETs might appear and also our spherical models, of differing sizes and intensities, used to simulate NETs. In addition, we describe our procedure for forming projection data using the Monte Carlo package SIMIND (Ljungberg and Strand, 1989). This section contains simulation details including system geometry, camera orbit, modeling of patient scatter and collimator interactions. In Sec. 7.2.2 we describe a family of five parallel hexagonal-hole collimators to be evaluated. Sec. 7.2.3 includes background information explaining our ideal observer including a mathematical statement of the ideal observer itself.

Applying the ideal observer yields a performance scalar figure of merit, ALROC, for each collimator. We use the ALROC to rank our collimator performances and to also assess the effects of signal size and search tolerance, i.e. the allowed radius about a true lesion within which an estimated location is deemed correct. Our experimental results are presented in Sec. 7.3.

In Sec. 7.4.1 we discuss the results and point out some limitations of our study. We have already briefly mentioned relevant previous work in this Introduction. However, with the aid of the technical vocabulary established in Secs 7.1-3, we discuss in much more detail in Sec. 7.4.2 the relation of our work to additional (beyond that already cited in this Introduction) work by others. We end the paper with a Conclusion in Sec. 7.5.

The aim of this study is to apply, using physically realistic simulations, an ideal observer to evaluate collimators for SPECT NET imaging. The observer carries out a novel and highly clinically relevant task of detecting the presence of and estimating the location of NET metastases in the liver.

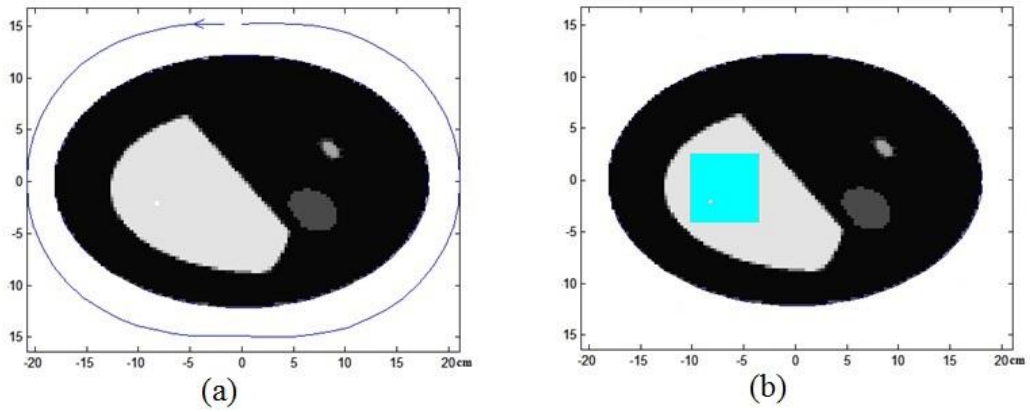

Figure 7.1 (a) One slice of the MCAT activity phantom. The phantom shows background tissue, liver, spleen and a small section of heart. Brighter regions have higher activity. The blue line indicates the contour-following collimator orbit used by the ECAM camera. **(b) The same slice but with a blue rectangle indicating the 21 x 21 search region superposed.** The bright dot shows a small signal at one of the possible locations in the search region.

7.2. Methods

7.2.1. Image Formation

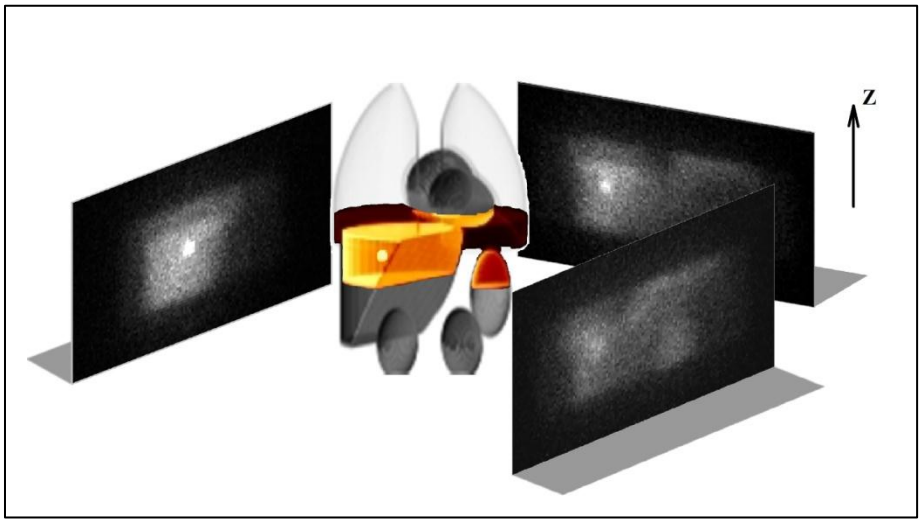

Figure 7.2 This summarizes our acquisition scheme. The Z direction (perpendicular to transaxial planes) is shown. Organs from the full MCAT activity phantom are shown with the colored region delineating our truncated (in Z) version. Camera faces at three angles are shown with camera-to-object distance increased for visual clarity. The collimator is not shown. A bright

dot in the object is our signal, and its projection can be seen in the camera faces. The signal intensity is exaggerated here for visual clarity, but the relative intensity of the camera faces is otherwise accurate.

We used SIMIND to generate Monte Carlo (MC) projection data to simulate SPECT NET imaging. As our background object, we used the MCAT activity and attenuation phantoms and set relative activities, appropriate for In^{111} pentetreotide, to heart : lung : liver : background : spleen : bone =19:5:20:1:16:1 (He et al., 2005). The activity and attenuation voxels were cubic $(3.12mm)^3$ and the camera pixels (bins) were $(3.12mm)^2$. Figure 7.1(a) shows one slice of the MCAT activity source phantom; each slice comprises 128 x 128 voxels. As shown in Figure 7.2 we used a truncated version of the 128-slice source MCAT limited to 32 transaxial object slices that included the liver. We embedded this in a coregistered attenuation phantom of 64 transaxial slices. The 32-slice activity phantom was symmetrically placed within the 64-slice attenuation phantom, so that there were 16 attenuation slices above the activity phantom and 16 attenuation slices below. The activity slice of Figure 7.1 (a) containing the search region is placed in the middle of the truncated activity phantom and is the 16th slice (as counted from the top, i.e. closest to the head) of the 32-slice phantom. Figure 7.2 shows the untruncated activity phantom (in shades of grey) and the truncated activity phantom (colored). Figure 7.1 (a) also shows the camera orbit and thus establishes the placement of the camera faces, 3 of which are shown in Figure 7.2. The camera faces were set to 128 x 64 bins corresponding to a size of 20cm in the Z-direction by 40cm laterally (Figure 7.2 defines the Z direction). Following protocols at our own clinic, data were acquired at 60 angles over 360^o.

Figure 7.1(b) shows the same slice as Figure 7.1(a) but with the search region (a blue rectangle) superposed on the liver. The search region comprises 21 x 21 pixels and a signal (NET) can be centered on any of the 441 locations with equal probability. The NET's, described below, are 3D spheres, but their centers are confined to the 2D search region.

The activity and attenuation phantoms were truncated in Z to limit the computational complexity of MC simulation. The camera face dimension was also limited in Z to track the phantom truncation. A major concern regarding these truncations is that photons originating in slices above or below the truncated slices might contribute significant septal penetration and

patient scatter counts to the projection data and affect observer performance. In the Discussion in Sec. 7.4.1, we show that this truncation did not affect our results.

The NETs ("signals") were simulated as spheres of varying diameters and contrasts. In any imaging trial, we could have a signal-absent case, or a signal-present case with a single signal present somewhere in the search region of Figure 7.1(b). We used signals of three diameters: 0.312 cm (1voxel), 0.936 cm (3 voxels) and 1.560 cm (5 voxels). Since the diameters corresponded to an odd number of voxels, each signal had a center voxel. We filled the voxels with relative activity in the following way: If the voxel was completely contained in a sphere of radius r (measured from the middle of the center voxel), the activity in the voxel was weighted by unity. If the voxel was partially occupied by the sphere, the weight was assigned as the percent of voxel volume that intersected the sphere. One exception was the small 1-voxel signal: we simply used a cubic voxel instead of trying to interpolate it onto a spherical volume. In our digital phantom, we define the *location* of the signal as that of the center voxel. In all cases, the center voxel of the spherical signal is located somewhere in the search region.

For the small (0.312 cm dia.), medium (0.936 cm dia.) and large (1.56 cm dia.) signals, we used signal contrasts of 28.66, 2.00, 0.43, respectively. We defined signal contrast as $\left(\textit{activity in the signal region - activity in the background region}\right)$ / $\left(\textit{activity in the background region}\right)$.

Contrasts were chosen to satisfy two constraints: (1) The contrast was chosen such that the scalar performance metric ALROC was not saturated at too high or too low a value. (2) The total integrated signal activity was the same for each diameter, thus smaller signals were bright and large signals dim. These constraints were needed to facilitate the observer study and were not meant to reflect clinically accurate tumor contrasts.

In running SIMIND, counts were summed from two ±10% energy windows centered at the 171 and 245 keV In^{111} lines. The effects of patient scatter (up to order 1) and attenuation were included along with collimator geometrical response, septal penetration and septal scatter. Backscatter effects and Pb X-ray generation were not modeled since these added negligible counts. For the ELEGP collimator, for instance, the total counts due to backscatter and Pb generation were only 9% of those due to patient scatter and 8% of those due to collimator scatter and penetration. A typical spectrum of the acquired data is shown in Figure 7.3. Note that the 171 keV window includes a considerable amount of downscatter from the 245 keV line. All MC

runs used an NaI(Tl) crystal of thickness 0.95 cm and ignored interactions with the collimator cover. The intrinsic camera resolution was fixed at 0.34 cm. Thus our Monte Carlo simulation captured the major physical effects associated with image formation.

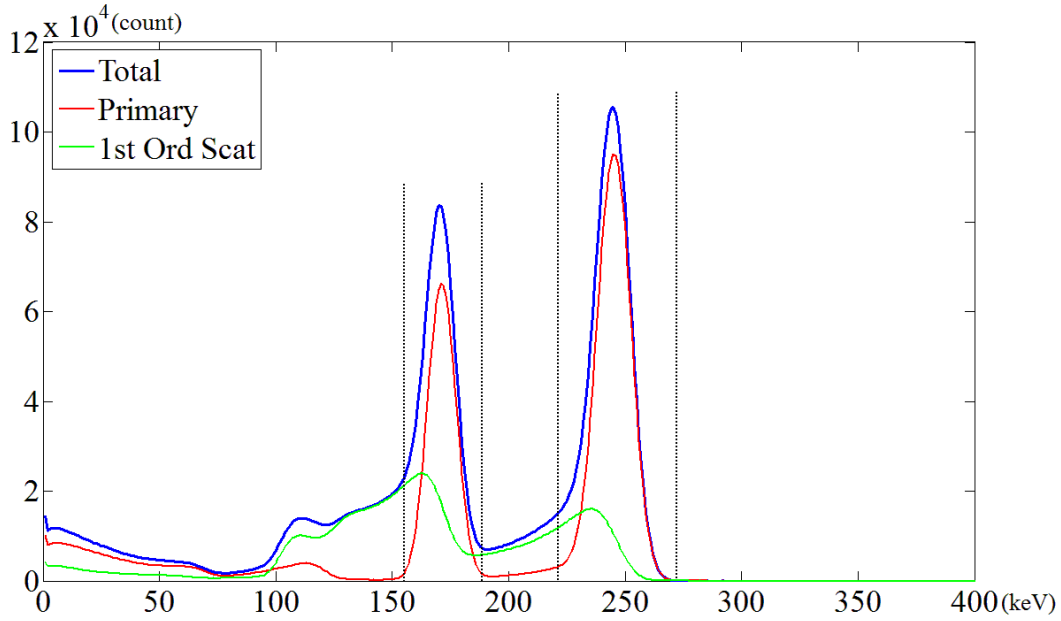

Figure 7. 3 Energy spectrum for a typical In-111 SPECT study showing the contributions of all photons and the contributions by single-scattered photons. Vertical dotted lines delineate the ±10% energy windows. This spectrum is obtained using the ELEGP collimator (described below).

A large enough number of source events was simulated to create essentially noiseless projection data. This data was scaled to reflect a clinical level of 20M counts/frame for a Siemens E-CAM medium energy (SEME) collimator and this scaled count level was further reduced to reflect the truncation of our object and our camera face. For other collimators, the 20M count figure was scaled by the efficiency relative to the E-CAM before undergoing the further reductions due to phantom thickness and camera dimensions. To the "noiseless" projections, we added Poisson noise. That is, noiseless projections were formed from enough source events that the Poisson noise added to the projection data was much larger than fluctuations in the MC "noiseless" data. The camera face images at three different angles in Figure 7.2 show the result of this procedure for the ELEGP collimator. In Figure 7.4, camera face images from one angle are shown for each of our 5 collimators.

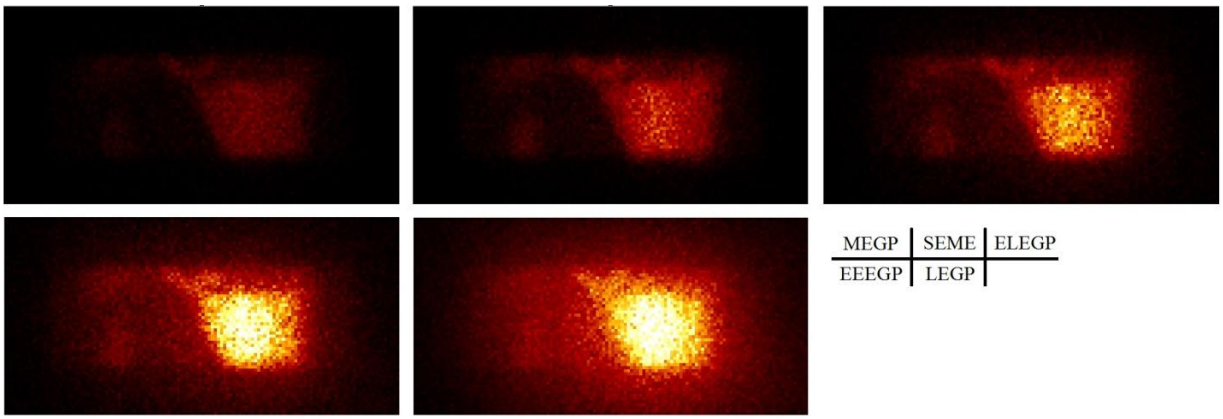

Figure 7.4 Camera face images (at one angle) of the signal-absent phantom for the five collimators. The color scale is common to the five images, except for the LEGP image which is saturated in the hot region. The relative efficiencies of the collimators are reflected by the brightness of the figures.

We can now write the image formation step mathematically. Let vector $\mathbf{b} = (b_n, n = 1,...,N)$ represent the background activities so that b_n is the activity (counts/voxel-sec) of voxel n. Note that $n = 1,...,N$ indexes all voxels in all slices. Let vector $\mathbf{s}^j$ be the spherical tumor signal centered at location j, where j is an element of the 21 x 21 search grid. Since $\mathbf{s}^j$ is a vector, it has components s_n^j indicating the activity of the signal in voxel n. The vector $\mathbf{g}$ has components g_m, $m = 1,...,M$ that are the noisy (with Poisson noise) count levels in the m th detector bin. Note that $m = 1,...,M$ indexes all bins at all angles, i.e. at all camera faces. The matrix $\mathbf{H}$, our system matrix, comprises elements H_{mn} that are proportional to the probability that a photon emitted from voxel n gets detected at bin m. Thus $g_m = [\mathbf{Hb}]_m + [\mathbf{n}]_m$ is the count level in bin m resulting from the forward projection of $\mathbf{b}$ into camera bin m. Here $[\mathbf{n}]_m$ is the Poisson noise in bin m. The forward projection operations $\mathbf{Hb}$ and $\mathbf{Hs}^j$ are not computed analytically, but are carried out by Monte Carlo simulations since such products model the effects of patient scatter and collimator effects.

For a signal-present object, we can thus summarize image formation as

$$\mathbf{g} = \mathbf{Hb} + \mathbf{Hs}^j + \mathbf{n} \tag{7.1}$$

where the zero-mean Poisson noise vector $\mathbf{n}$ is written without loss of generality as an additive term. (For a signal-absent object, the term $\mathbf{Hs}^j$ in (7.1) is deleted.) Note that the $\mathbf{H}$ in (7.1) depends on collimator choice, and we now give the collimator details.

$$R_g = \frac{D}{L}(L + a)$$

where D is bore diameter, L the bore length and a =16 cm is a reference distance between a point source and the surface of the collimator. Note that R_g is for a geometric psf (point spread function) only and ignores contributions to the psf from septal scatter and penetration. We also computed R_c, the resolution distance of the psf when collimator scatter and septal penetration are included. To obtain R_c, we performed MC, including collimator interactions, using a point source in air at 16 cm from the surface of the collimator. The distance R_c is a measure of the width of the resulting psf observed on the camera face. Since this psf has long tails, we took R_c equal to the full width at 1/30th max of the psf. The psf, a star-like pattern, is anisotropic, but we measured R_c in a direction perpendicular to a collimator septum (equivalent to measuring the psf width along the longest "arms" of the star-like psf). The Relative Total Sensitivity is measured with an MC run using our MCAT signal-absent phantom and is the ratio of the total counts received in the energy windows relative to the counts received in the SEME reference collimator for the same effective exposure time. In table 1, the Scat/Pri Ratio is the total number of patient scattered photons (Scat) in the energy windows divided by the number of photons (Pri) in the energy windows that did not interact with phantom. The Pen/Tot Ratio is defined as the number of photons in the energy windows that penetrated the collimator septa divided by the total number of photons accepted into the energy windows. The Coll Scat/Tot Ratio definition is the same except the numerator counts photons scattered from the septa. The last line of Table 1, Relative Geometrical Sensitivity, lists collimator sensitivity due to geometrical effects only, i.e. efficiency due to photons excluding those that penetrate through or scatter from the collimator. Note that this sensitivity is about the same for all collimators (except the MEGP), showing that the rise in Total Sensitivity is due in part to the contribution from collimator effects. A further insight into the collimator behaviors can be seen in Figure 7.6, which plots collimator psf profiles. The profiles are taken in a direction perpendicular to the septa.

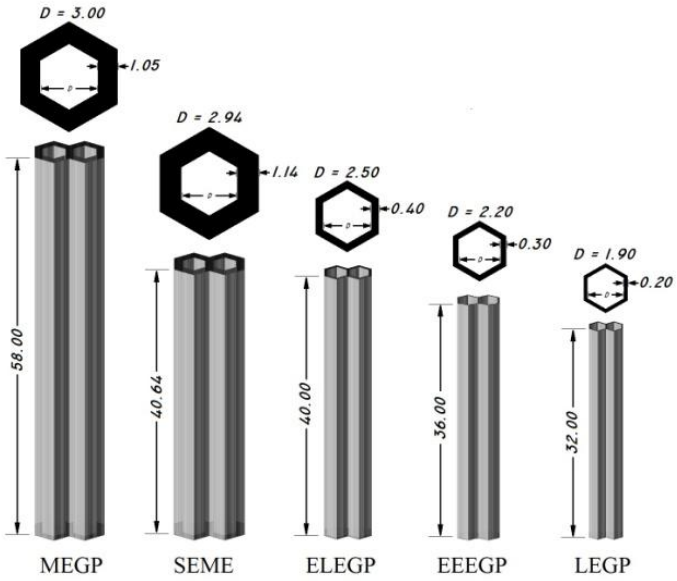

Figure 7.5 The 5 hexagonal collimators evaluated in our study. The cross sections and bore lengths are drawn to scale. The bore lengths, diameters and septal thickness distances are shown. All units are in mm. The definition of "Diameter" D is indicated in the cross-sectional diagrams at the top of each collimator. MEGP is "Medium Energy General Purpose", SEME is "Siemens E-CAM Medium Energy", ELEGP is "Extended Low Energy General Purpose", LEGP is "Low Energy General Purpose", and EEEGP is "Extended ELEGP".

Table 7.1 Properties of the Five Collimators. See Text for Extensive

Parameters	MEGP	SEME	ELEGP	EEEGP	LEGP
Diam(mm)	3.00	2.94	2.50	2.20	1.90
Septa(mm)	1.05	1.14	0.40	0.30	0.20
Length(mm)	58.00	40.64	40.00	36.00	32.00
R_g (mm) at 16 cm	11.84	15.24	13.13	12.58	11.97
R_c (mm) at 16 cm	24.96	29.24	30.02	38.22	173.94
Scat/Pri Ratio	43.41%	42.74%	33.08%	27.94%	25.13%
Relative Tot Sens	0.519	1.000	1.681	2.752	6.747
Pen/Tot Ratio	4.79%	7.33%	28.19%	44.92%	63.01%
Coll Scat/Tot Ratio	2.80%	3.22%	14.18%	19.86%	22.53%
Relative Geo Sens	0.480	0.895	0.969	0.969	0.976

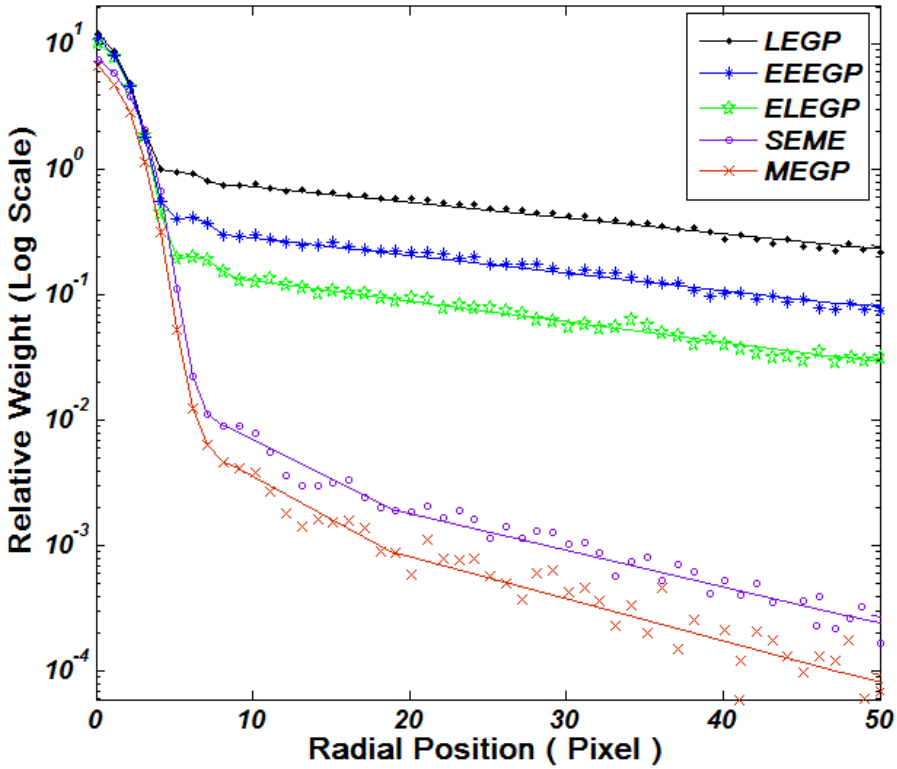

Figure 7. 6 Collimator point spread functions (psf) obtained from an MC simulation of a point source in air at 16 cm from the collimator face. The central core of each psf relects geometrical resolution and the long tails reflect collimator penetration and collimator scatter. The heights of the curve reflects the relative efficiencies of the 5 collimators.

The five collimators are designed to exhibit an increase in total sensitivity while keeping geometrical sensitivity ("Relative Geo/Sens" in table 1) constant. (One exception is MEGP which has a lower geometrical sensitivity than the other collimators) The increase in total sensitivity (as we move from left to right in the row labeled "Relative Total Sensitivity" in table 1) can thus be ascribed to collimator scatter and penetration. Indeed the values for "Pen/Tot Ratio" and "Coll Scat/Tot Ratio" increase as we move from left to right in Table 7.1. As a consequence of these design edicts, the amount of patient scatter, as measured by "Scat/Pri Ratio" decreases as we move from left to right in Table 7.1. Thus the camera data **g** include, within the two energy windows, all combinations of photons that experienced or did not experience patient scatter, collimator scatter and penetration.

7.2.3 Ideal Observer and performance evaluation

We summarize the essential features of our ideal observer with enough detail to support the material in other sections of the paper and to allow replication of our numerical experiments. The reader is referred to our previous work in (Khurd and Gindi, 2005a) and (Zhou et al., 2008) for more mathematical detail.

To help explain our ideal observer, we first consider the simpler task of detecting a signal in a known location (no localization) and review relevant aspects of the widely known ROC (receiver operating characteristic) curve that is used in performance analysis of this task. To explain the ROC computational steps we follow the widely cited tutorial article by Metz (Metz, 1978)

Consider the detection problem of deciding whether a known signal at a known location is present or absent. The decision strategy involves applying an observer, which may be a numerical observer, to the data (in our case $\mathbf{g}$) to generate a scalar observer response $t(\mathbf{g})$. The quantity $t(\mathbf{g})$ is roughly a "confidence" that the signal is present. The response $t(\mathbf{g})$ is compared to a decision threshold τ and the signal is deemed present if $t(\mathbf{g}) \geq \tau$ and absent if $t(\mathbf{g}) < \tau$. This threshold comparison allows the observer to choose hypotheses $\mathcal{H}_1$ (signal present) or $\mathcal{H}_0$ (signal absent).

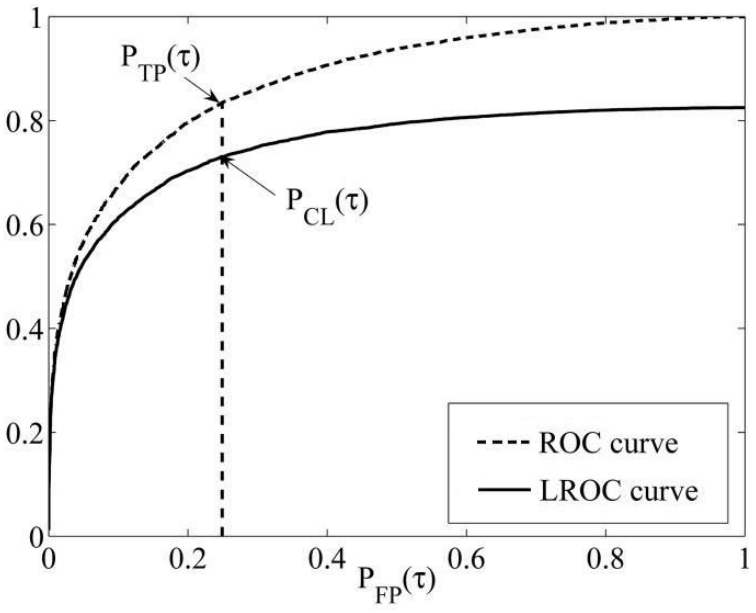

Figure 7.7 Typical ROC and LROC curves. The ordinate is $P_{TP}(\tau)$ for ROC and $P_{CL}(\tau)$ for LROC. For both curves the abscissa is $P_{TP}(\tau)$**.** See text for detailed explanation.

The performance of the observer as a function of decision threshold can be tracked by the ROC curve. The ROC curve, shown in Figure 7.7, plots the true positive rate $P_{TP}(\tau)$, which is the probability of deciding signal present when it is indeed present, versus the false positive rate $P_{FP}(\tau)$ which is the probability of deciding signal present when it is absent. As τ varies, the ROC curve is generated. The ROC curve can be generated by first accumulating a sample set of observer responses $t_+^k, \quad k=1,...,P_{samp}$ obtained by applying the observer to P_{samp} signal-present images, and a sample set of observer responses $t_-^k, \quad k=1,...,Q_{samp}$ obtained by applying the observer to Q_{samp} signal-absent images. (The subscripts "+" and "-" denote signal-present and signal-absent, respectively.) Normalized histograms of t_+ and t_- are integrated from τ to $+\infty$ to obtain $P_{TP}(\tau)$ and $P_{FP}(\tau)$, respectively, and the ROC curve is then plotted. Note that this method of computing the ROC curve is appropriate for a numerical observer, but for human observers, who report a limited number of "integer confidence ratings" (e.g. an integer from 1 to 5), elaborate fitting procedures and parametric forms (binomial) for the ROC curve are typically

used to derive and evaluate the ROC curve. For our case, the histogramming approach is appropriate.

Note that the height of the ROC curve ($P_{TP}(\tau)$ at a given $P_{FP}(\tau)$) is a threshold-dependent performance measure of the observer. A widely accepted threshold *independent* performance measure is the *area* under the ROC curve, AROC. One can compute AROC by numerical integration.

So far we have not specified the form of the observer and our description applies to any observer $t(\mathbf{g})$. (Here we interchangeably use "$t(\mathbf{g})$" to mean observer or observer response.) The *ideal* observer has the property that it yields the *maximum* AROC amongst all observers (Barrett and Myers, 2003). It turns out that for the ROC case, the ideal observer is the well-known likelihood ratio $\mathrm{LR}(\mathbf{g})$:

$$t(\mathbf{g}) = \mathrm{LR}(\mathbf{g}) = \frac{p(\mathbf{g} \mid \mathcal{H}_1)}{p(\mathbf{g} \mid \mathcal{H}_0)} = \frac{p(\mathbf{g} \mid \mathbf{b}+\mathbf{s})}{p(\mathbf{g} \mid \mathbf{b})}$$

where $p(\mathbf{g} \mid \mathcal{H}_1)$ and $p(\mathbf{g} \mid \mathcal{H}_0)$ are the data likelihoods under the hypotheses $\mathcal{H}_1$ and $\mathcal{H}_0$, and $\mathbf{g}$ and $\mathbf{b}$ are as defined earlier. The vector $\mathbf{s}$ is a known signal at a known location.

We now use our description of the ROC curve and its associated performance measure AROC to help explain the LROC (localization ROC) curve (Swensson, 1996) in Figure 7.7 and its associated performance measure ALROC suitable for our more complex detection/localization task. We again use a threshold-based decision strategy: an observer examines $\mathbf{g}$ and yields $t(\mathbf{g})$. The observer *also* yields $l(\mathbf{g})$ which is an estimate of the location of the signal. If $t(\mathbf{g}) < \tau$, the signal is deemed absent and hypothesis $\mathcal{H}_0$ is selected. If $t(\mathbf{g}) \geq \tau$, then the signal is deemed present at $l(\mathbf{g})$ and (for l discrete) hypothesis $\mathcal{H}_l$, indicating signal present at l, is chosen. Note that $l = 1, ..., L$ indexes the search region.

The observer performance as a function of τ can be tracked by the LROC curve shown in Figure 7.7. The abscissa is $P_{FP}(\tau)$, the probability of deciding $\mathcal{H}_l$ when there is no signal. The ordinate is $P_{CL}(\tau)$, the "correct localization rate". A signal is deemed correctly localized if the signal is actually present, hypothesis $\mathcal{H}_l$ is chosen, and the estimated signal location $l(\mathbf{g})$ is sufficiently close to the true location. By "sufficiently close" we mean that the true signal

location is within a circular tolerance region $T(l(\mathbf{g}))$ about $l(\mathbf{g})$. We shall use the term TOL to indicate the radius of T. As τ varies, the LROC curve is generated.

The LROC curve can be generated by using P_{samp} samples of signal-present images and Q_{samp} samples of signal-absent images. The observer is applied to the Q_{samp} signal-absent images to derive responses $(t_-^k(\mathbf{g}), l_-^k(\mathbf{g}))$, $k = 1,...,Q_{samp}$. The observer is also applied to the P_{samp} signal-present images to obtain $(t_+^k(\mathbf{g}), l_+^k(\mathbf{g}))$, $k = 1,...,P_{samp}$. Amongst the P_{samp} observer pairs, we retain only those that are correctly localized (i.e. $l_+^k(\mathbf{g})$ is within the tolerance region). In our experiments we used $P_{samp} = 4000$ and $Q_{samp} = 2000$. Suitably normalized histograms of t_+ and t_- are again obtained and integrated from τ to $+\infty$ to obtain $P_{CL}(\tau)$ and $P_{FP}(\tau)$, respectively.

As for the case of AROC, the area under the LROC curve, ALROC, is used as a performance metric for the observer. Note that the LROC curve, and hence ALROC, will depend on one's choice of TOL, the tolerance radius. This choice is sometimes made (Paik et al., 2004) according to clinical criteria.

The ideal observer for the LROC case makes use of a generalized likelihood ratio

$$\mathrm{LR}(\mathbf{g}, \mathbf{s}^j) \equiv \frac{p(\mathbf{g} \mid \mathcal{H}_j)}{p(\mathbf{g} \mid \mathcal{H}_0)}$$

We also need to know $p(\mathbf{s}^j)$, the prior probability that the signal will be found at j. With these definitions, the ideal observer (Khurd and Gindi, 2005a) becomes

$$t(\mathbf{g}) = \max_{l \in \{1,...,L\}} \sum_{j \in T(l)} p(\mathbf{s}^j) \mathrm{LR}(\mathbf{g}, \mathbf{s}^j) \tag{7.2}$$

$$l(\mathbf{g}) = \arg \max_{l \in \{1,...,L\}} \sum_{j \in T(l)} p(\mathbf{s}^j) \mathrm{LR}(\mathbf{g}, \mathbf{s}^j) \tag{7.3}$$

$$\textit{Decide } H_{l(\mathbf{g})} \textit{ if } t(\mathbf{g}) > \tau, \textit{ else decide } H_0 \tag{7.4}$$

We can describe the action of this ideal observer as follows: for a given $\mathbf{g}$ the ideal observer computes a "likelihood ratio image" $\mathrm{LR}(\mathbf{g}, \mathbf{s}^j)$ as a function of spatial positions j within the search region. The $\mathrm{LR}(\mathbf{g}, \mathbf{s}^j)$ is weighted by the prevalence $p(\mathbf{s}^j)$, which in the work here will be taken as uniform without loss of generality. This "image" is convolved with the binary circular tolerance mask $T(l)$. Finally the maximum value of this "image" is taken as $t(\mathbf{g})$ and the location of the max as $l(\mathbf{g})$. Given the $(t(\mathbf{g}), l(\mathbf{g}))$ response, the appropriate hypothesis is chosen

as described earlier, or, equivalently, as summarized in Eq (7.4). A rough (not mathematically exact) description of the action of the ideal observer in (7.2)-(7.4) is this: It examines the likelihood ratios at all locations in the search region and delivers the value and location of the largest one.

The one quantity still needed an expression for $\mathrm{LR}(\mathbf{g},\mathbf{s}^j)$ for SPECT. From the Poisson nature of the data noise in SPECT, it is easy to show that

$$\mathrm{LR}(\mathbf{g},\mathbf{s}^j) = \prod_m \exp(-[\mathbf{H}\mathbf{s}^j]_m)\left(1+\frac{[\mathbf{H}\mathbf{s}^j]_m}{[\mathbf{H}\mathbf{b}]_m}\right)^{g_m} \tag{7.5}$$

where m indexes detector bins. In (7.5) $\mathbf{g}$ (and its components g_m) are the noisy camera bin values for one noise realization (one image) and $\mathbf{s}^j$ indicates that LR is evaluated under a hypothesis of the signal location at j. The products $\mathbf{H}\mathbf{s}^j$ and $\mathbf{H}\mathbf{b}$ are precomputed by SIMIND and include effects of geometrical response, patient scatter, septal penetration and septal scatter. Once again, we emphasize that $\mathrm{LR}(\mathbf{g},\mathbf{s}^j)$, the key part of the ideal observer, is evaluated in the detector space $\mathbf{g}$ but used to make inferences about the signal location in the object (MCAT) space.

Note that the ideal observer is applied to the unmodified raw data **g**. No postprocessing or patient scatter correction is applied to **g**. The reason for this is that (Barrett and Myers, 2003) the ideal observer has a perfect model of (as seen in Eq.(7.5)) and compensates for all effects in **g**. Any postprocessing of **g** that does not make use of some sort of new information not contained in $\mathrm{LR}(\mathbf{g},\mathbf{s}^j)$ will not affect the performance of the ideal observer.

7.3. Results

Table 7.2 and its graphical representation Figure 7.8 summarize our main results. For each signal diameter, ALROC vs. collimator is evaluated. This is done for four tolerances TOL=1,2,3,4 where TOL is measured in units of pixel width. For a given tolerance, the ALROC rank order is independent of signal size. The rank order is: ELEGP best, then SEME, EEEGP, MEGP and the worst is LEGP. This rank order is violated in only one instance, for SEME and EEEGP at TOL=1, where the rank order is flipped. For a given signal size, this same rank order is generally preserved for all tolerances. The exception again involves SEME and EEEGP, where as

tolerance increases, the rank order of SEME overtakes that of EEEGP. That ALROC falls with tolerance as expected, since as tolerance drops, the probability of correct localization necessarily falls.

The values in Table 7.2, each a single sample of ALROC, lack errors bars because these values are essentially deterministic. We can verify this by applying a t-test to each pair of collimators in one row of Table 7.2, i.e. a pair of collimators that share the same signal size and tolerance. We applied a two-sided t-test using N=25 ALROC samples. The choice of t-test is consistent with sample size, N=25. (The t-test also applies for a more subtle reason. There is no correlation between ALROC samples across collimator classes due to the fact that there is a single "patient", the MCAT. The measurement noise is the only source of uncertainty.) Typical p-values are $p = 1.39\times10^{-11}$ for the SEME vs. ELEGP pair at TOL=4 and signal size (diam) =15.60mm. The highest p-value is $p= 2.65\times10^{-8}$ for the SEME vs. EEEGP at TOL=2 and signal size = 9.36mm. The p-values are small because the standard deviation of the ALROC samples for either collimator of a pair is very small compared to the difference of means. For example, for the SEME and EEEGP collimators pair at TOL=2 and signal size = 9.36 mm, we have *(difference of means)/ (avg standard deviation)* = 2.2. This set yields a p-value of 2.65×10^{-8}. That the standard deviations are so small is due to the large number of observer responses, i.e. 4000 signal-present and 2000 signal-absent, used to obtain each ALROC. Observer responses are cheap to compute, hence the use of 4000 and 2000 observer responses.

The best collimator, ELEGP, has a moderate efficiency that is due in part to allowing some amount of collimator penetration and collimator scatter. If we increase this collimator penetration and collimator scatter further, as with the EEEGP and LEGP, performance drops. If, on the other hand, these collimator contributions are reduced as in the MEGP and SEME, performance again drops. So for this work, a moderate amount of collimator penetration and collimator scatter appears beneficial.

For this limited set of five collimators, we tried designs that increased total sensitivity while keeping geometrical sensitivity fairly constant. Thus the increase in total sensitivity can be ascribed to the septal scatter and penetration. We thus once again conclude that, qualitatively, a controlled amount of septal scatter and penetration in the collimator design is beneficial for our task. The five collimator designs included two clinically used collimators (SEME and MEGP), a simulated collimator (ELEGP proposed in (Mähler et al., 2012)), another simulated collimator

(EEEGP) that we designed as a more liberal (more penetration and scatter) version of ELEGP, and an extreme (very high penetration and scatter) collimator, the LEGP (used in (Mähler et al., 2012)).

Table 7. 2 ALROC Results for Different Collimators at Different Tolerance Radii TOL (in units of pixel width) and Different Signal Sizes

Signal Diam	TOL Radii	MEGP	SEME	ELEGP	EEEGP	LEGP
3.12 mm	TOL = 1	0.3755	0.3854	0.4586	0.4232	0.3129
	TOL = 2	0.5814	0.6325	0.6757	0.6071	0.447
	TOL = 3	0.7189	0.7873	0.8122	0.745	0.5854
	TOL = 4	0.7845	0.8669	0.8791	0.8177	0.6979

Signal Diam	TOL Radii	MEGP	SEME	ELEGP	EEEGP	LEGP
9.36 mm	TOL = 1	0.2945	0.3223	0.4262	0.3717	0.2205
	TOL = 2	0.4904	0.5481	0.6217	0.5420	0.3488
	TOL = 3	0.6434	0.7325	0.7749	0.6862	0.5158
	TOL = 4	0.7249	0.8236	0.8446	0.7781	0.6428

Signal Diam	TOL Radii	MEGP	SEME	ELEGP	EEEGP	LEGP
15.60 mm	TOL = 1	0.2069	0.2694	0.3344	0.2879	0.1669
	TOL = 2	0.3769	0.4774	0.5193	0.4391	0.2793
	TOL = 3	0.5387	0.6585	0.6775	0.5985	0.4305
	TOL = 4	0.6517	0.7714	0.7822	0.7125	0.5736

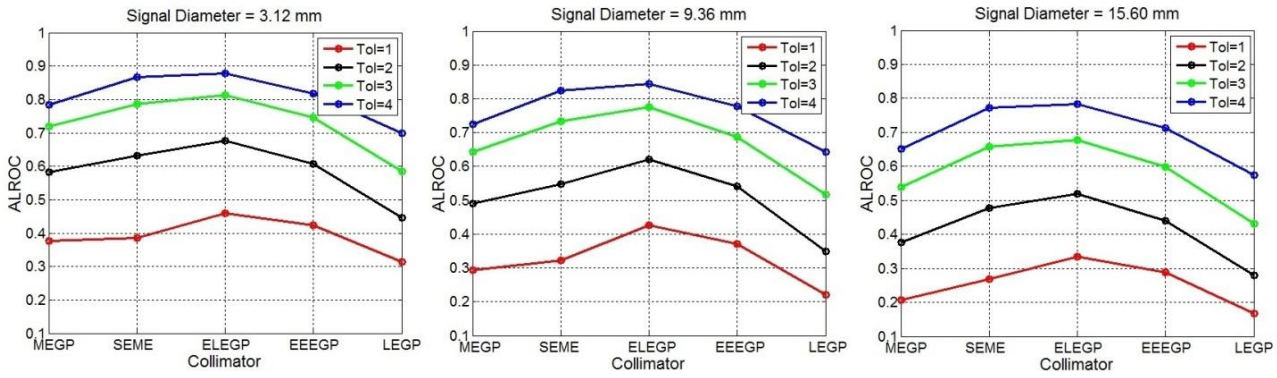

Figure 7.8 Graphical representation of the information in Table 7.2. Each graph is for a different signal diameter, and plots ALROC vs. collimator choice. Each graph plots these quantities for four tolerance distances, TOL=1,2,3,4 in units of pixel width.

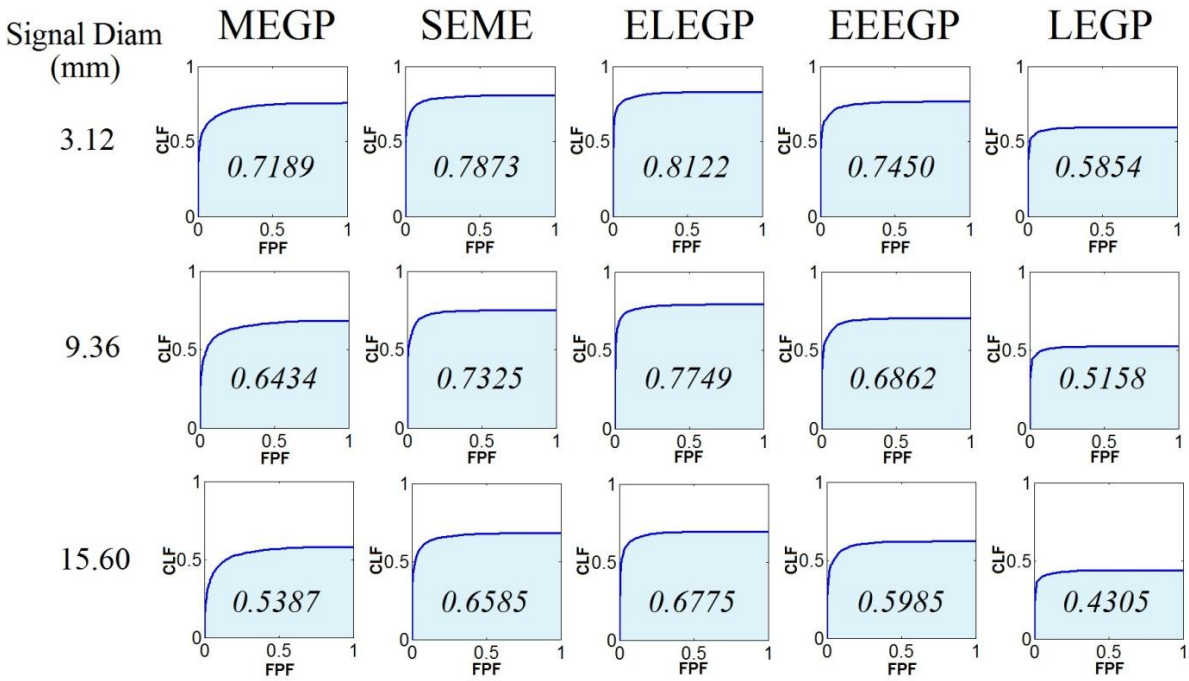

Figure 7.9 The LROC curves at one tolerance (TOL=3). The axes of each plot, CLF = "Correct Localization Fraction" and FPF = "False Positive Fraction", are the same as P_{CL} and P_{FP}, respectively. (The acronyms were used for visual clarity.) The number in the shaded portion is the area ALROC.

Figure 7.9 shows the 15 LROC curves for the 5 collimators at the three signal sizes and for TOL=3. Qualitatively, all 60 LROC curves exhibit the steep rise followed by long plateau as seen in Figure 7.9. We observed the following interesting feature of our set of LROC curves: Given any two LROC curves, the one with the larger ALROC will lie entirely above the one with the smaller ALROC. That is, the curves do not cross. A minor exception to the "curves do not cross" rule occurs for some pairs of MEGP and LEGP LROC curves, where the curves can cross at very small (<0.05) values of P_{FP}. As stated in (Metz, 1978), the choice of operating point, i.e. choice of a (P_{FP}, P_{CL}) pair on the ALROC curve, is governed by one's estimate of costs and benefits of correct and incorrect decisions. That the ALROC curves do not cross means that one can choose a collimator based only on ALROC ranking and disregard any consideration of choice of operating point.

7.4. Discussion

Our main contributions are the introduction of a new, clinically relevant detection/localization task for collimator evaluation, and the conclusions reached with this methodology for collimator performance tradeoffs involving septal penetration and collimator scatter for the multi-energy In^{111} pentetreotide agent. In Sec. 7.4.1 we discuss limitations of the study. In Sec. 7.4.2 we discuss work related to ours by virtue of using the ideal observer on the raw data for system optimization, as well as some relevant related work on aperture optimization and evaluation. (Some of this work was briefly mentioned in the Introduction in Sec. 7.1 but, given the concepts and vocabulary we established in Sec. 7.2 and Sec. 7.3, we can discuss this related work in more detail.)

7.4.1 Limitations of the study

The study did not model backscatter effects and Pb X-rays, and the justification for this has been given in Sec. 7.2.1. We limited patient scatter simulation to only first order, thus saving considerable computation time. The higher-order patient scatter would have added only a small amount (< 20%) of extra scatter into the energy windows.

In Sec. 7.2.1 we indicated that the activity and attenuation phantoms were truncated in the Z direction and that a thicker phantom might contribute additional counts that could affect ALROC. To verify that our truncated phantom was adequate, we used the LEGP collimator to image phantoms of varying thicknesses. One criterion to compute an acceptable maximal phantom thickness is to compute, using the point spread functions of Figure 7.6, the contribution to the camera from photons emanating from the top (highest-Z) slice. We would then take an acceptable phantom thickness as the one where a source point in the middle of this "top" slice contributes a zero or negligibly small number of counts to the camera face at any angle. We chose the LEGP collimator because it has the worst-case point spread function as seen in Figure 7.6. However, we did not use the above method to establish phantom thickness. Instead, we used a different task-based criterion. We varied the phantom thickness and the camera dimension in Z and computed ALROC. This ALROC is affected by geometrical effects, collimator scatter and penetration and also patient scatter from source points emanating from the extreme top and bottom slices. We observed that as phantom thickness increased, ALROC decreased until it reached a stable value. We chose our final phantom thickness to correspond to the point where ALROC stabilized. The choice was conservative in that collimators other than the LEGP reached this critical point at smaller thicknesses. We tested three phantom thicknesses: 5cm (activity)/10cm (attenuation), 10cm/20cm and 20cm/40cm. (The full MCAT phantom was 40 cm thick.) The corresponding camera face dimensions in Z were 10cm, 20cm and 40 cm. Camera face widths were fixed at 40 cm. The ALROCs for the 10cm/20cm and 20cm/40cm phantoms were the same, but the 5cm/10cm phantom led to a higher ALROC. For this reason we used the 10cm/20cm phantom. Thus, with the truncated phantom, we benefit from reduced computational requirements while being assured that ALROC values are unaffected.

Our results showed that the rank order of performance for our five collimators was barely affected by signal size. However, our collimator family was chosen to test tradeoffs due to collimator scatter and penetration and patient scatter. A more natural family of collimators to test signal size would be one where only geometric resolution R (and the related geometric efficiency) increased. In fact, in our previous work (Zhou and Gindi, 2009) we used such a family of collimators. We found that the optimal R increased with signal size. However, this study had the many limitations discussed previously in Sec. 7.1.

Our study did not include the effects of background variability (Barrett and Myers, 2003) (He et al., 2008) (Lee et al., 2013) (Faris et al., 1999). Background variability is the statistical variation of background uptake in the population of patients seen by the imaging systems. This form of variability can affect task performance. Our own previous work (Zhou and Gindi, 2009) had limitations as discussed in Sec. 7.1, but did include a variability model that captured small-scale "textural" fluctuations in **b** with a correlated Gaussian process. As we increased variability by raising the amplitude of this form of noise, ALROC decreased. Computational complexity precluded the inclusion of background variability in the current work, but remains a topic for future effort.

7.4.2 Related work

As mentioned in Sec. 7.1, a study of three collimators (MEGP, ELEGP, LEGP) in the context of In^{111} pentetreotide NET visualization was done in (Mähler et al., 2012). Because (Mähler et al., 2012) evaluated performance in terms of contrast and CNR (contrast-to-noise ratio) of various sized spherical NET signals in the *reconstructed* domain, the performance was not only dependent on collimator choice, but was also dependent on the parameters and compensation methods of the reconstruction algorithm. However, our collimator ranking was independent of postprocessing of the projection data, unlike the case in (Mähler et al., 2012). Interestingly, the rank order of performance of these 3 collimators was the same that we obtained (albeit we evaluated two additional collimators.) That is, the ranking in (Mähler et al., 2012) was (from best to worst) ELEGP, MEGP, LEGP while ours was ELEGP, SEME, EEEGP, MEGP, LEGP.

Moore (Moore et al., 1998) studied a large family of 35 lead square-hole collimators in the context of simulated In^{111} *planar* imaging of a simple water phantom containing a lesion of variable but known size at a fixed distance (16 cm) from the collimator. For a detection-only task, the optimal collimator had dimensions (diam, septum, length) = (2.47mm, 0.67mm, 34.8mm) which for an equivalent hexagonal collimator (Moore et al., 2005) is (2.65mm, 0.74mm, 37.4mm). For this experiment, the optimal collimator exhibited much less septal penetration than that of the ELEGP in our SPECT experiment. In a related study (Moore et al., 2005), using the multi-energetic Ga^{67} isotope, 20 parallel-square-hole lead collimators, and planar imaging of a phantom similar to that in (Moore et al., 1998), collimators were again evaluated by detection

performance. The optimal collimator had better performance than two commercially used medium-energy collimators.

Others have used the ideal observer or approximate ideal observer applied to the raw camera data $\mathbf{g}$ in SPECT for purposes of system optimization. Lee (Lee et al., 2013) used a scanning linear observer for the task of detecting, localizing and estimating the activity in cardiac perfusion defects for the purpose of evaluating the pinhole arrangement in a clinical pinhole-base cardiac SPECT scanner. Other related work can be found in (Ghaly et al., 2013) (He et al., 2008) (Hesterman et al., 2005) (Zeng and Gullberg, 2002) (Kamphuis et al., 1999) and (Westcott et al., 2007). Our own relevant previous work (Zhou and Gindi, 2009) was discussed in Sec. 7.1.

7.5. Conclusion

We have described a new approach to collimator evaluation. This approach relies on evaluating the raw camera data with a numerical ideal observer. The task carried out by the observer is detection *and localization* of the tumor. The collimator evaluations were carried out in the context of In^{111} pentetreotide in neuroendocrine tumor imaging.

We evaluated five collimators. These five collimators were chosen to isolate the effects of collimator scatter and penetration by keeping the geometrical sensitivity of each collimator nearly the same while allowing a controlled increase in counts due to collimator scatter and penetration. We found that the ELEGP performed best as measured by ALROC, the area under the LROC curve. The rank order of performance amongst the five collimators was independent of the three signal sizes tested. For a given signal size, this same rank order is almost always preserved for all tolerances. The ELEGP allowed a moderate amount of septal scatter and penetration. Other collimators in our family allowed much less or much more septal scatter and penetration and resulted in reduced performance. Thus building a controlled amount of collimator scatter and penetration into the collimator design improves performance on our SPECT task.

Chapter 8 An Approach to System Optimization for X-Ray Photon-Counting Systems Using Performance on a Detection/Localization Task

This chapter is based on work originally published as (Lu et al., 2013). The work was primarily my own.

8.1 Introduction

In previous chapters, we performed imaging system optimization for a detection/localization task for SPECT. For transmission (X-ray) systems, we previously mentioned that it was difficult to write the exact likelihood for the raw data because of the complexity of modeling the detector and electronics. Recent developments in photon counting X-ray detectors have led to detectors fast enough to count individual X-ray photons and measure their energy, as is done for γ-rays for the lower count rate SPECT. This development in the X-ray field opens the door to study system optimization for a transmission system. The system modeled here is a computed tomography (CT) scanner, not a DBT system.

CT is an effective modality in lung nodule detection and surveillance. In recent years, there has been a growing interest in the use of CT for lung nodule screening applications (Sone et al., 1998) (Henschke, 2000). In the early-stage lung cancer diagnosis, both detection of small nodules and monitoring the nodule growth are very important (Piyavisetpat et al., 2005) (Paul et al., 2007). In order to facilitate regular check-up of a patient, it is important to perform low-dose CT (Yankelevitz et al., 1999). However, a low-dose CT scan will significantly degrade the image quality due to noise and artifacts (Hsieh, 2009). Numerous reconstruction methods have been proposed (Liu, 2014) (Tian et al., 2011) (Shieh et al., 2015) to reduce such effects. Such reconstruction methods are often iterative methods with regularization (the OSTR algorithm in chapter 4 can be modified to be a such algorithm for CT). In order to be effective, the choice of reconstruction parameters is very crucial (e.g. iteration number, value of the regularizer β and

edge preserver parameter δ in OSTR). Different combinations of parameters can result in very different image quality (Liu, 2014). Unfortunately, usually there isn't a best/automatic way rather than empiricism to select a best parameter combination, which is one of the reasons that iterative reconstruction methods haven't been widely adopted in the clinics. To evaluate/optimize the reconstruction images, physical measures of detector performance or simple measures of image quality (such as contrast-to-noise ratio) are often used. A more rational way to optimize medical imaging system acquisition parameters is to maximize a scalar performance metric for a clinically relevant task (Barrett and Myers, 2003). In the lung cancer screening application, one task is to determine if a lesion (signal) is present in a region of the image, and if so, to estimate its location in the region.

One approach to task-based system optimization is to use a numerical ideal observer applied to the raw projection data (Park et al., 2009, Lu et al., 2014c). This approach ignores the details of the reconstruction algorithm and acknowledges that all potentially useful information is contained in the projection data. Other approaches (Das et al., 2011) have used performance of the detection/localization task applied to the reconstruction - not sinogram - using a human emulating numerical observer in place of the ideal observer.

Our main contribution is to formulate the ideal observer for our task that is applicable to any photon-counting X-ray tomographic system. We demonstrate this ideal observer using a simulation of 2D CT system in which the task is to localize and detect a small lung nodule. Performance on this task is used to optimize angular schedule of projection acquisitions at a fixed dose. We also verify that performance increases as dose is increased.

Photon-counting X-ray systems have received much attention in recent years. Due to the development of electronics and X-ray sensors, it is now possible to design detectors that count individual X-ray photons in high-flux clinical beam with a good energy resolution (Bellazzini et al., 2015). Note that X-ray photon-counting detection in itself is a vast topic and discussion of such is beyond the scope of this thesis. In this study, we assume the CT system is equipped with an idealized photon-counting detector which assumes the detector can count each individual photon with an infinite energy resolution. The fact that photon-counting systems can deliver the energy of a given photon can be exploited for many purposes (Schlomka et al., 2008). Also, they

are applicable for imaging at much lower doses (Barber et al., 2009) than those used by conventional X-ray energy integrating detectors.

The topic of X-ray photon-counting detectors in medical imaging is current and vast, and detailed discussion is beyond our scope. For our purposes, we assume the existence of an X-ray detector not limited by count rate (i.e. counts individual photons) and with perfect energy resolution and perfect spatial resolution limited only by detector size.

8.2 Ideal Observer - Monoenergetic Case

8.2.1. Imaging equation

We first consider the monoenergetic case. It will be useful to first define some quantities so that we can state the imaging equation. Let vector $\boldsymbol{\mu}$ be the attenuation object, $\boldsymbol{\mu}_{s_j}$ be the attenuation of a signal s_j located at pixel j and $\boldsymbol{\mu}_b$ be the background. Then $\boldsymbol{\mu} = \boldsymbol{\mu}_b + \boldsymbol{\mu}_{s_j}$ if a signal is present. Let $\mathbf{A}$ be the system matrix defined so that $\mathbf{A}\boldsymbol{\mu}$ is a forward projection. Unlike conventional systems with energy-integrating detectors, for photon-counting systems, the photon counts (denoted by vector $\mathbf{I}$ with components $[\mathbf{I}]_m$) are Poisson distributed. In this case, the monoenergetic imaging equation is $\mathbf{I} \sim Poisson\left[\mathbf{b}\exp\left(-\mathbf{A}\boldsymbol{\mu}\right)\right]$, where $\mathbf{b}$ is the air (blank) scan. Note that this equation is the same as eq.(2.4) except that the noise is actually Poisson, not approximately Poisson. Here we do not consider scatter.

8.2.2. Ideal observer

The original formulation of the mathematical ideal observer, independent of modality specific details, was done by our group in (Khurd and Gindi, 2005a). In Chapters 6 and 7, we discussed the details about such ideal observer. An observer is ideal in the sense that it maximizes ALROC, the area under the LROC curve. The LROC curve, which plots the probability of correct detection and localization vs the false positive rate, is an extension of the ROC curve, which plots the probability of correct detection vs the false positive rate. The areas under curves are useful performance measures. In Chapter 7 and in (Lu et al., 2014c) that

Chapter 7 is based on, we specialized the ideal observer to emission imaging and used it for collimator optimization in SPECT.

Here we offer a new formulation for the ideal observer applicable to photon counting X-ray imaging. The ideal observer is formulated as follows: Let $T(l)$ be a circular tolerance region about pixel l within which a detected signal is said to be correctly localized. In this case the ideal-observer strategy is

$$\begin{aligned} t(\mathbf{I}) &= \max_{l\in\{1,\ldots,L\}} \sum_{j\in T(l)} p(\boldsymbol{\mu}_{s_j})\mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j}) \\ l(\mathbf{I}) &= \arg\max_{l\in\{1,\ldots,L\}} \sum_{j\in T(l)} p(\boldsymbol{\mu}_{s_j})\mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j}) \\ &\textit{Decide } \mathcal{H}_{l(\mathbf{I})} \textit{ if } t(\mathbf{I}) > \tau, \textit{ else decide } \mathcal{H}_0 \end{aligned} \tag{8.1}$$

where $p(\boldsymbol{\mu}_{s_j})$ is the prior probability of a signal at j, and the generalized likelihood ratio LR is defined below. $\mathcal{H}_0$ is the hypothesis of no signal present, $\mathcal{H}_{l(\mathbf{I})}$ the hypothesis that a signal is present at l, and τ is a decision threshold. The ideal observer delivers a response $t(\mathbf{I})$ and a location estimate $l(\mathbf{I})$. From $t(\mathbf{I})$ and $l(\mathbf{I})$, ALROC can be calculated. One may notice that eq.(8.1) is identical to eq. (7.2)-(7.4) except that here we use $\mathbf{I}$ instead of $\mathbf{g}$ and $\boldsymbol{\mu}_{s_j}$ instead of $\mathbf{s}_j$.

8.2.3.Likelihood Ratio

The term $\mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j})$ is a generalized likelihood ratio defined as

$$\mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j}) \equiv \frac{p\left(\mathbf{I}\,|\,\mathcal{H}_j\right)}{p\left(\mathbf{I}\,|\,\mathcal{H}_0\right)} \tag{8.2}$$

where $\mathcal{H}_j$ is the *hypothesis* that a signal present and located at location j and $\mathcal{H}_0$ represents the signal absent hypothesis. Let $\bar{\mathbf{g}}_b \equiv \mathbf{A}\boldsymbol{\mu}_b$ and $\bar{\mathbf{g}}_{s_j} \equiv \mathbf{A}\boldsymbol{\mu}_{s_j}$, where $\bar{\mathbf{g}}_{s_j}$ and $\bar{\mathbf{g}}_b$ are the noiseless projection data (line-integrals) of $\boldsymbol{\mu}_b$ and $\boldsymbol{\mu}_{s_j}$. Then eq.(8.2) can be rewritten as

$$\mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j}) \equiv \frac{p\left(\mathbf{I} \mid \bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}\right)}{p\left(\mathbf{I} \mid \bar{\mathbf{g}}_b\right)} \tag{8.3}$$

Here we give the derivation of $p\left(\mathbf{I} \mid \bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}\right)$ and $p\left(\mathbf{I} \mid \bar{\mathbf{g}}_b\right)$: For a given signal present on a background $\boldsymbol{\mu}_b + \boldsymbol{\mu}_{s_j}$, which is a signal-present case, the probability of detector bin *m* receiving $[\mathbf{I}]_m$ photons is

$$p([\mathbf{I}]_m \mid \bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}) = \frac{\exp\left\{-[\mathbf{b}]_m \exp\left\{-[\bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j})]_m\right\}\right\}\left\{[\mathbf{b}]_m \exp\left\{-[\bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}]_m\right\}\right\}^{[\mathbf{I}]_m}}{[\mathbf{I}]_m!} \tag{8.4}$$

which follows because $[\mathbf{I}]_m \sim \mathit{Poisson}\left[[\mathbf{b}]_m \exp\left([-\mathbf{A}\boldsymbol{\mu}]_m\right)\right]$, $[\bar{\mathbf{g}}]_m = [\mathbf{A}\boldsymbol{\mu}]_m$ and the counts for an individual detector are Poisson distributed. Note that the number of photons leaving the X-ray tube and headed toward a detector (i.e. along ray *m*) is Poisson with mean $[\mathbf{b}]_m$. Due to attenuation, we have a Bernoulli selection process with $p = e^{-[\mathbf{A}\boldsymbol{\mu}]_m}$. Thus we get eq.(8.4). Then joint probability of all detector bins becomes:

$$p(\mathbf{I} \mid \bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}) = \prod_m \frac{\exp\left\{-[\mathbf{b}]_m \exp\left\{-[\bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j})]_m\right\}\right\}\left\{[\mathbf{b}]_m \exp\left\{-[\bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}]_m\right\}\right\}^{[\mathbf{I}]_m}}{[\mathbf{I}]_m!} \tag{8.5}$$

Eq.(8.5) follows from the Poisson independence argument as given for SPECT. Here, *m* indexes all rays (combinations of detectors and X-ray source positions). The counts for a given ray are Poisson and independent from those of other rays. For a signal absent case, joint probability of all detector bins follows directly as:

$$p(\mathbf{I} \mid \bar{\mathbf{g}}_b) = \prod_m \frac{\exp\left\{-[\mathbf{b}]_m \exp\left\{-[\bar{\mathbf{g}}_b)]_m\right\}\right\}\left\{[\mathbf{b}]_m \exp\left\{-[\bar{\mathbf{g}}_b]_m\right\}\right\}^{[\mathbf{I}]_m}}{[\mathbf{I}]_m!} \tag{8.6}$$

Then we insert eq.(8.5) and (8.6) into eq.(8.3), and after some algebra, obtain:

$$\begin{aligned} \mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j}) = \exp&\left\{-\sum_m \left(\left[\bar{\mathbf{g}}_{s_j}\right]_m [\mathbf{I}]_m\right)\right\} \\ &\cdot\exp\left\{\sum_m \left([\mathbf{b}]_m \exp\left(-\left[\bar{\mathbf{g}}_b\right]_m\right)\right)\right\} \\ &\cdot\exp\left\{-\sum_m \left([\mathbf{b}]_m \exp\left(-\left[\bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}\right]_m\right)\right)\right\} \end{aligned} \tag{8.7}$$

Eq.(8.7) is one of the contributions of this work. Note that we cannot take a *log* in eq.(8.7) due to the sum in eq.(8.1). Eq (8.7) can be evaluated with numerical stability despite its strange form. Note that for this section we assumed the monochromatic case. For a polychromatic case, we discuss it in Sec.8.4.

8.3 CT Simulations

We performed an optimization on a simulated 2-D CT version of a system modeled on the Siemens SOMOTOM scanner, but assumed a photon counting capability. For the purposes of this simulation we assumed a monoenergetic case. We performed our simulation using a one-slice 512 x 512 MCAT phantom. The voxel size is 0.78125 mm. We chose a slice intersecting the heart as one can see in Figure 8.1. The human ribs can be seen as the white segments. The light grey "ring" shaped part represents spine and the dark grey area is the lung. The X-ray energy is simulated as 120 keV. The attenuation coefficients were assigned according to (ICRU, 1989). The task was to localize a small lung nodule (3mm in diameter) in a lung search region (2.1cm x 2.1cm) as shown in the 512 x 512 phantom in Figure 8.1. The search region is unrealistically small due to the computational complexity of evaluating eq.(8.7) at all locations in the search region.

The system optimization goal was to decide the best (i.e. max ALROC) number of angular acquisitions given a fixed dose (in total an average of 525,000 photons is received). For low-dose CT, various strategies are used when lowering the dose. Since CT is used for screening purposes, we want as low a dose as possible. We can turn down the mAs to a fraction of that used in a diagnostic dose. At the very low dose, it may turn out that the signal is more easily seen (Liu, 2014) by portioning the dose to a few angles. But if it is too few, performance goes down and if too many angles, performance goes down. So we seek the optimum number of equiangular acquisitions to maximize performance.

Since we simulated 2-D CT, the projection data is a sinogram. The detector contains 672 bins. Figure 8.2 shows 2D 84 angles sinogram superposed with an artificially enhanced sine wave due to the signal.

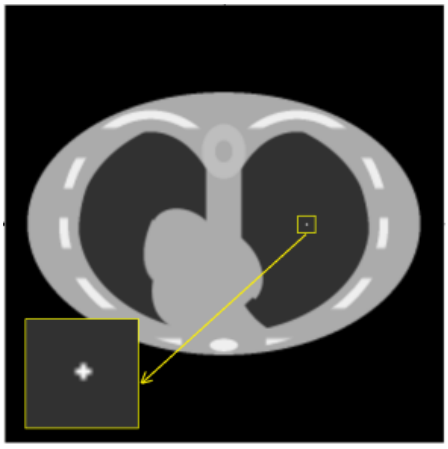

Figure 8.1 MCAT phantom and simulated signal

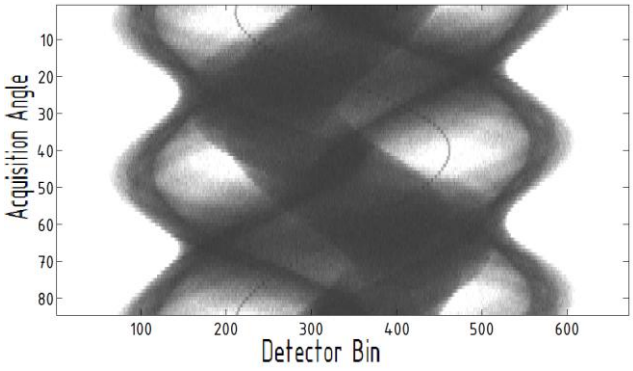

Figure 8.2 Sinogram including an artificially enhanced signal

To acquire the raw data $\mathbf{I}$ for 7, 14, 28, 35, 42, 56, 84 and 168 angles, we first simulated the noiseless raw data at a calculated dose level for each case. Since the total number of counts for each case is same, the mean angular dose for each case are 75,000, 37,500, 18,750, 15,000, 12,500, 9,375 and 3,125 photons. I added Poisson noise through MATLAB to the mean counts to obtain noisy realizations. Note for each case, we generated 2000 signal present cases and 2000 signal absent cases. We then calculated the LR for all the noisy realizations through eq.(8.7).

Note that the LR requires knowledge of blank scan $\mathbf{b}$, signal-alone noiseless line-integrals $\bar{\mathbf{g}}_{s_j}$ and background noiseless line-integrals $\bar{\mathbf{g}}_b$, thus we also used the simulator to obtain them. After calculating the LR, we applied ideal observer in eq.(8.1) to evaluate each different angle portioning strategy and finally get the ALROC scalar for different cases. Figure 8.3 plots ALROC vs. (equispaced) angle number, showing that 42 angles is best. Note the tolerance radius used here is 1.5 pixel, which results in that the tolerance region is the same as the signal size.

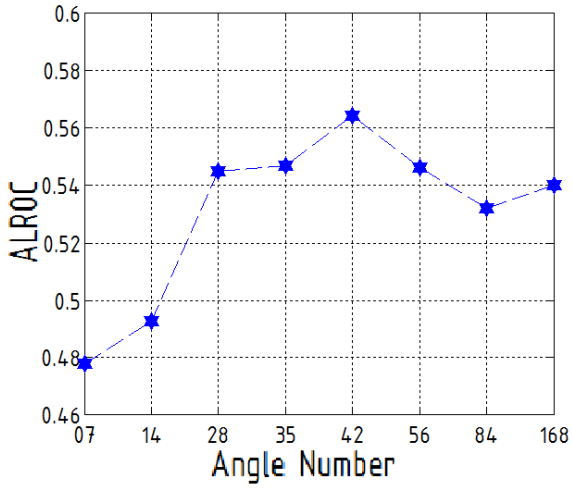

Figure 8.3 ALROC at different angle numbers and fixed dose for lung nodule searching. Error bar is too small to show.

We performed a second experiment in which the goal was to track ALROC as dose (total counts) increased. The angle number was fixed at 105. As seen in Figure 8.4, the ALROC rises with dose, as expected, until it saturates at a value of unity, corresponding to perfect detection and localization of the lung nodule. The curve shows the dose needed to attain a desired performance level. If the task were different, say simple detection with ALROC replaced by AROC, the area under the ROC curve, the performance vs dose curve would have the same trend, but would be quantitatively different and saturate at a different dose. To the extent that the detection/localization task is deemed more clinically realistic than a simple detection task where the location of the lesion is fixed (after all, the radiologist has to search the region), the ALROC curve is more relevant.

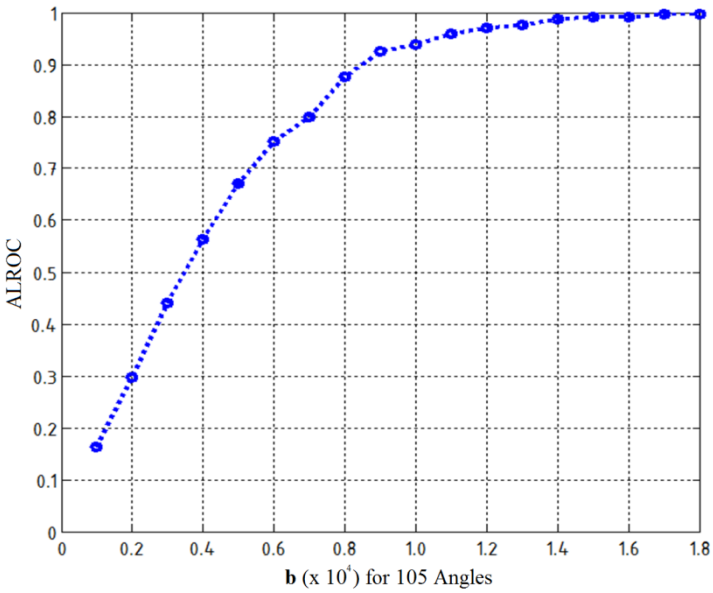

Figure 8.4 ALROC at different air scan counts for 105 angles

8.4 Extension to Poly-energetic case

We have shown the specific likelihood ratio form in eq.(8.2) for a monochromatic X-ray source and the derivation. Here we make an extension for a polychromatic X-ray source. For a given energy *E1*, a given signal present on a background becomes $\boldsymbol{\mu}_b^{E1}+\boldsymbol{\mu}_{s_j}^{E1}$, which is a signal present case. The probability of detector bin *m* receiving $[\mathbf{I}^{E1}]_m$ photons is

$$p([\mathbf{I}^{E1}]_m \mid \bar{\mathbf{g}}_b^{E1}+\bar{\mathbf{g}}_{s_j}^{E1}) = \frac{\exp\left\{-[\mathbf{b}^{E1}]_m \exp\left\{-[\bar{\mathbf{g}}_b^{E1}+\bar{\mathbf{g}}_{s_j}^{E1})]_m\right\}\right\}\left\{[\mathbf{b}^{E1}]_m \exp\left\{-[\bar{\mathbf{g}}_b^{E1}+\bar{\mathbf{g}}_{s_j}^{E1}]_m\right\}\right\}^{[\mathbf{I}^{E1}]_m}}{[\mathbf{I}^{E1}]_m!} \tag{8.8}$$

For another energy *E2*, we have

$$p([\mathbf{I}^{E2}]_m \mid \bar{\mathbf{g}}_b^{E2}+\bar{\mathbf{g}}_{s_j}^{E2}) = \frac{\exp\left\{-[\mathbf{b}^{E2}]_m \exp\left\{-[\bar{\mathbf{g}}_b^{E2}+\bar{\mathbf{g}}_{s_j}^{E2})]_m\right\}\right\}\left\{[\mathbf{b}^{E2}]_m \exp\left\{-[\bar{\mathbf{g}}_b^{E2}+\bar{\mathbf{g}}_{s_j}^{E2}]_m\right\}\right\}^{[\mathbf{I}^{E2}]_m}}{[\mathbf{I}^{E2}]_m!} \tag{8.9}$$

Since a sum of two independent Poisson distribution is still Poisson with the means adding (Wang, 1997), if the X-ray source simultaneously emits *E1* and *E2*, then the probability of detector bin *m* receiving $[\mathbf{I}]_m$ photons become

$$p([\mathbf{I}]_m \mid \bar{\mathbf{g}}_b^{E1} + \bar{\mathbf{g}}_{s_j}^{E1}, \bar{\mathbf{g}}_b^{E2} + \bar{\mathbf{g}}_{s_j}^{E2}) = \frac{e^{-[\mathbf{b}^{E1}]_m \exp\left\{-[\bar{\mathbf{g}}_b^{E1} + \bar{\mathbf{g}}_{s_j}^{E1})]_m\right\} - [\mathbf{b}^{E2}]_m \exp\left\{-[\bar{\mathbf{g}}_b^{E2} + \bar{\mathbf{g}}_{s_j}^{E2})]_m\right\}} \left\{[\mathbf{b}^{E1}]_m e^{-[\bar{\mathbf{g}}_b^{E1} + \bar{\mathbf{g}}_{s_j}^{E1}]_m} + [\mathbf{b}^{E2}]_m e^{-[\bar{\mathbf{g}}_b^{E2} + \bar{\mathbf{g}}_{s_j}^{E2}]_m}\right\}^{[\mathbf{I}]_m}}{[\mathbf{I}]_m !} \tag{8.10}$$

If we generalize eq.(8.10) for all the energies, then we have for bin *m*

$$p([\mathbf{I}]_m \mid \bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}) = \frac{\prod_E \left\{\exp\left\{-[\mathbf{b}^E]_m \exp\left\{-[\bar{\mathbf{g}}_b^E + \bar{\mathbf{g}}_{s_j}^E)]_m\right\}\right\}\right\} \cdot \left\{\sum_E \left\{[\mathbf{b}^E]_m \exp\left\{-[\bar{\mathbf{g}}_b^E + \bar{\mathbf{g}}_{s_j}^E]_m\right\}\right\}\right\}^{[\mathbf{I}]_m}}{[\mathbf{I}]_m !} \tag{8.11}$$

Note that the *E* superscript on $\bar{\mathbf{g}}$ is now deleted since *E* is now an index used to sum/multiply all energies and $\bar{\mathbf{g}}$ is assumed polyenergetic.

Using the previously discussed independence of Poisson counts, the joint probability of signal present for all detector bins becomes:

$$p(\mathbf{I} \mid \bar{\mathbf{g}}_b + \bar{\mathbf{g}}_{s_j}) = \frac{\prod_m \left[\prod_E \left\{\exp\left\{-[\mathbf{b}^E]_m \exp\left\{-[\bar{\mathbf{g}}_b^E + \bar{\mathbf{g}}_{s_j}^E)]_m\right\}\right\}\right\} \cdot \left\{\sum_E \left\{[\mathbf{b}^E]_m \exp\left\{-[\bar{\mathbf{g}}_b^E + \bar{\mathbf{g}}_{s_j}^E]_m\right\}\right\}\right\}^{[\mathbf{I}]_m}\right]}{\prod_m \{[\mathbf{I}]_m !\}} \tag{8.12}$$

The joint probability of signal absent follows directly:

$$p(\mathbf{I} \mid \bar{\mathbf{g}}_b) = \frac{\prod_m \left[\prod_E \left\{\exp\left\{-[\mathbf{b}^E]_m \exp\left\{-[\bar{\mathbf{g}}_b^E]_m\right\}\right\}\right\} \cdot \left\{\sum_E \left\{[\mathbf{b}^E]_m \exp\left\{-[\bar{\mathbf{g}}_b^E]_m\right\}\right\}\right\}^{[\mathbf{I}]_m}\right]}{\prod_m \{[\mathbf{I}]_m !\}} \tag{8.13}$$

Finally, by plugging (8.12) and (8.13) into (8.3), the generalized likelihood ratio for a polychromatic X-ray becomes:

$$\begin{aligned}\mathrm{LR}(\mathbf{I},\boldsymbol{\mu}_{s_j}) &= \frac{p(\mathbf{I}\,|\,\bar{\mathbf{g}}_b+\bar{\mathbf{g}}_{s_j})}{p(\mathbf{I}\,|\,\bar{\mathbf{g}}_b)} \\ &= \prod_m \left(\left[\frac{\sum_E \left\{ [\mathbf{b}^E]_m \exp\left\{ -[\bar{\mathbf{g}}_b^E + \bar{\mathbf{g}}_{s_j}^E]_m \right\} \right\}}{\sum_E \left\{ [\mathbf{b}^E]_m \exp\left\{ -[\bar{\mathbf{g}}_b^E]_m \right\} \right\}} \right]^{[\mathbf{I}]_m} \right) \\ &\cdot \exp\left\{ \sum_E \left(\sum_m [\mathbf{b}^E]_m \exp\left\{ -[\bar{\mathbf{g}}_b^E]_m \right\} \right) \right\} \\ &\cdot \exp\left\{ -\sum_E \left(\sum_m [\mathbf{b}^E]_m \exp\left\{ -[\bar{\mathbf{g}}_b^E + \bar{\mathbf{g}}_{s_j}^E]_m \right\} \right) \right\}\end{aligned} \tag{8.14}$$

Again, *m* indexes the cross product of all X-ray tube positions and all detectors, that is, *m* indexes all rays. For the poly-energetic case, we have made use of the photon-*counting* capability of modern detectors but *not* their energy *discrimination* capability. The measured data $[\mathbf{I}]_m$ is integrated over all energies. To carry out the ALROC experiment, we do need to know energy dependent quantities $\mathbf{b}^E$ and $\bar{\mathbf{g}}^E = \mathbf{A}\boldsymbol{\mu}^E$ as seen in (8.14). These quantities are integrated out via the $\sum_E$ operation. The energy discriminating aspect of these modern detectors can be exploited in many ways, but this is beyond the scope of this thesis.

Since real X-ray sources in medical systems are polychromatic, one would have to use the polyenergetic form (8.14) to conduct a more realistic ALROC experiment.

In this chapter, I derived the generalized likelihood form for X-ray transmission modality for both monoenergetic and polychromatic cases. An idealized 2-D X-ray CT simulation showed the feasibility of applying the ideal observer of detection/localization to a lung nodule medical scenario.

Chapter 9 Summary and Future Work

I have worked on acquisition and reconstruction strategies for the new modality CE-DE-DBT, and on optimizing acquisition (collimator characteristics) for In-111 SPECT. Here, I present some problems one would need to address to extend this work.

For CE-DE-DBT I worked on the advantages of RS vs. SR acquisition/reconstruction, the assessment of different reconstruction algorithms, and on scatter correction (SC), I simulated interleaved acquisition by culling projection data sets in software. Actually implementing interleave entails technical problems such as rapid filter switching. We implicitly model X-ray firing as if it were step-and-shoot mode, but in most scanners it would be continuous tube motions, so we would have to model focal spot motion blur.

The interleaved RS acquisition mode is designed to minimize patient motion artifacts. Yet we assumed no motion (the phantom was motionless) and instead checked RS vs. SR for residual artifacts in the reconstruction. Clearly, some experiment with controlled phantom motion is needed to firm up the RS results.

Regarding the reconstruction algorithms for the RS/SR project, all of them have many free parameters. As in any medical reconstruction project some rational means, perhaps task-oriented, would be needed to select these. For OSTR in particular, the results were anecdotal with parameters β and δ selected by hand to achieve qualitatively good looking reconstructions of the small iodine inserts.

My SC method has the advantage of rapid implementation and yields good results for phantom testing. One item to further address includes the prevention of detector saturation outside the breast shadow since this would destroy data in the wing region needed for interpolation. We did experiments simulating the CC view only; the MLO view could be done but it requires much more work in interpolating along trajectories not parallel to the x-axis. Though we had only one set of patient data (mammo not DBT), the entire project would need be extended to the clinical setting should such data become available.

For the SPECT project, we tested only 5 collimator designs. But the parameter space of collimator design (bore shape, length, width, septal thickness) is large and one really needs to test a much larger family of collimator designs to be confident in conclusions.

Our philosophy was that ALROC delivered by the ideal observer viewing the projection data **g** was a stringent enough criterion for collimator selection. One might argue that the reconstruction algorithm and the collimator design should be optimized jointly, with a human emulating model observer replacing the IO.

Bibliography

Anger, H. O. 1958. Scintillation Camera. *Review of Scientific Instruments,* 29, 27-33.

Barber, W. C., Nygard, E., Iwanczyk, J. S., Zhang, M., Frey, E. C., Tsui, B. M. W., Wessel, J. C., Wawrzyniak, G., Hartsough, N. E., Handhi, T. & Taguchi, K. Characterization of A Novel Photon Counting Detector for Clinical CT: Count Rate, Energy Resolution, and Noise Performance. SPIE Medical Imaging, 2009.

Barrett, H. H. & Myers, K. J. 2003. *Foundations of Image Science,* New York, Wiley Interscience.

Bellazzini, R., Breza, A., Spandrea, G., Minutia, M., Pinchera, M., Delogua, P., Ruvoa, P. L. D. & Vincenzib, A. 2015. PIXIE III: a very large area photon-counting CMOS pixel ASIC for sharp X-ray spectral imaging. *Journal of Instrumentation,* 10.

Berry, D., Cronin, K., Plevritis, S., Fryback, D., Clarke, L., Zelen, M., Mandelblatt, J., Yakovlev, A., Habbema, J., Feuer, E., Cancer, I. & Surveillance Modeling Network, C. 2005. Effect of screening and adjuvant therapy on mortality from breast cancer. *The New England journal of medicine,* 353, 1784-1792.

Beutel, J., Kundel, H. L. & Van Metter, R. L. 2000. Handbook of Medical Imaging, volume 1: Physics and Psychophysics. *Spie Press Bellingham, WA*.

Beyer, T., Townsend, D. W., Brun, T., Kinahan, P. E., Charron, M., Roddy, R., Jerin, J., Young, J., Byars, L. & Nutt, R. 2000. A combined PET/CT scanner for clinical oncology. *J Nucl Med,* 41, 1369-79.

Boone, J. M. & Cooper Iii, V. N. 2000. Scatter/primary in mammography: Monte Carlo validation. *Med. Phys.,* 27, 1818-31.

Boone, J. M., Nelson, T. R., Lindfors, K. K. & Seibert, J. A. 2001. Dedicated breast CT: radiation dose and image quality evaluation. *Radiology,* 221, 657-67.

Broeders, M. J., Ten Voorde, M., Veldkamp, W. J., Van Engen, R. E., Van Landsveld-Verhoeven, C., T Jong-Gunneman, M. N., De Win, J., Greve, K. D., Paap, E. & Den Heeten, G. J. 2015. Comparison of a flexible versus a rigid breast compression paddle:

pain experience, projected breast area, radiation dose and technical image quality. *Eur Radiol,* 25, 821-9.

Cadiot, G., Bonnaud, G., Lebtahi, R., Sarda, L., Ruszniewski, P., Le Guludec, D. & Mignon, M. 1997. Usefulness of Somatostatin Receptor Scintigraphy in the Management of Patients with Zollinger–Ellison Syndrome. Groupe de recherche et d'Etude du syndrome de zollinger-ellison (GRESZE). *Gut,* 41, 107-114.

Carton, A.-K. K., Ullberg, C., Lindman, K., Acciavatti, R., Francke, T. & Maidment, A. D. 2010. Optimization of a dual-energy contrast-enhanced technique for a photon-counting digital breast tomosynthesis system: I. A theoretical model. *Medical Physics,* 37, 5896-5907.

Carton, A. K., Li, J., Albert, M., Chen, S. & Maidment, A. 2006. Quantification for contrast-enhanced digital breast tomosynthesis. *SPIE Medical Imaging,* 6142, .

Cdc. 2010. *Cancer Among Women* [Online]. Available: http://www.cdc.gov/cancer/dcpc/data/women.htm.

Chen, L., Lu, Y., Hu, Y., Zhao, W. & Gindi, G. 2013. Impact of subtraction and reconstruction strategies on dual-energy contrast enhanced breast tomosynthesis with interleaved acquisition. *SPIE Medical Imaging,* 8668

Chen, X., Nishikawa, R., Chan, S.-T., Lau, A. B., Zhang, L. & Mou, X. 2012. Algorithmic scatter correction in dual-energy digital mammography for calcification imaging. *SPIE Medical Imaging,* 8313, .

Cherry, S. R., Sorenson, J. A. & Phelps, M. E. 2012. Physics in nuclear medicine. *Elsevier Health Sciences*.

Chien, A. A. & Karamcheti, V. 2013. Moore's Law: The First Ending and a New Beginning. *Computer,* 46, 48-53.

Cirs. *Tissue-equivalent phantom for mammography model 011A* *http://www.cirsinc.com/products/all/6/tissue-equivalent-phantom-for-mammography/?details=specs* [Online].

Das, M., Gifford, H. C., O'connor, J. M. & Glick, S. J. 2011. Penalized Maximum Likelihood Reconstruction for Improved Microcalcification Detection in Breast Tomosynthesis. *IEEE Trans. Med. Imaging.,* 30, 904-914.

Diaz, O., Dance, D. R., Young, K. C., Elangovan, P., Bakic, P. & Wells, K. 2014. Estimation of scattered radiation in digital breast tomosynthesis. *Physics in Medicine and Biology*, 59, 4375-90.

Dromain, C., Balleyguier, C., Muller, S., Mathieu, M., Rochard, F., Opolon, P. & Sigal, R. 2006. Evaluation of tumor angiogenesis of breast carcinoma using contrast-enhanced digital mammography. *AJR. American journal of roentgenology*, 187, W528-537.

Dromain, C., Thibault, F., Diekmann, F., Fallenberg, E. M., Jong, R. A., Koomen, M., Hendrick, R., Tardivon, A. & Toledano, A. 2012. Dual-energy contrast-enhanced digital mammography: initial clinical results of a multireader, multicase study. *Breast Cancer Research*, 14, .

Erdogan, H. & Fessler, J. 1999. Ordered subsets algorithms for transmission tomography. *Phys. Med. Biol.*, 44, 2835-51.

Faris, S., Wilson, D., Barrett, H. H., Dougherty, D., Gindi, G. & Hsiao, I.-T. 1999. Using a Digital Anatomical Phantom to Optimize an Imaging System. *SPIE Medical Imaging*.

Feng, S., D'orsi, C., Newell, M., Seidel, R., Patel, B. & Sechopoulos, I. 2014. X-ray scatter correction in breast tomosynthesis with a precomputed scatter map library. *Medical Physics*, 41, 319121-7.

Feng, S. & Sechopoulos, I. 2011. A software-based x-ray scatter correction method for breast tomosynthesis. *Medical Physics*, 38, 6643-6653.

Fessler, J. 2000. Statistical Image Reconstruction Methods for Transmission Tomography. *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis.* SPIE.

Gaemperli, O., Schepis, T., Valenta, I., Husmann, L., Scheffel, H., Duerst, V., Eberli, F. R., Luscher, T. F., Alkadhi, H. & Kaufmann, P. A. 2007. Cardiac image fusion from stand-alone SPECT and CT: clinical experience. *J Nucl Med*, 48, 696-703.

Gavenonis, S., Lau, K., Karunamuni, R., Zhang, Y., Ren, B., Ruth, C. & Maidment, A. 2012. Initial Experience with Dual-Energy Contrast-Enhanced Digital Breast Tomosynthesis in the Characterization of Breast Cancer. *IWDM 2012, LNCS vol 7361*, 32-39.

Ghaly, M., Links, J. M., Du, Y. & Frey, E. C. 2013. Model Mismatch and the Ideal Observer in SPECT. *SPIE Medical Imaging*, 86730K-86730K.

Gifford, H. C., King, M. A., Pretorius, P. H. & Wells, R. G. 2005. A Comparison of Human and Model Observers in Multislice LROC Studies. *IEEE Trans. Med. Imaging*, 24, 160-169.

Gunter, D. L. 2004. Collimator Design for Nuclear Medicine. *In:* WERNICK, M. N. & AARSVOLD, J. (eds.) *Emission Tomography-The Fundamentals of PET and SPECT.*

He, B., Du, Y., Song, X., Segars, P. & Frey, E. C. 2005. A Monte Carlo and Physical Phantom Evaluation of Quantitative In-111 SPECT. *Phys. Med. Biol.,* 50, 4169-4185.

He, X., Caffo, B. S. & Frey, E. C. 2008. Toward Realistic and Practical Ideal Observer (IO) Estimation for the Optimization of Medical Imaging Systems. *IEEE Trans. Med. Imaging,* 27, 1535-1543.

Henschke, C. I. 2000. Early lung cancer action project - Overall design and findings from baseline screening. *Cancer,* 89, 2474-2482.

Hesterman, J. Y., Kupinski, M. A., Furenlid, L. R. & Wilson, D. 2005. Experimental Task-Based Optimization of A Four-Camera Variable-Pinhole Small-Animal SPECT Systems. *Proc. of SPIE, Medical Imaging.*

Hill, M., Mainprize, J., Carton, A. K., Muller, S., Ebrahimi, M., Jong, R. A., Dromain, C. & Yaffe, M. J. 2013. Anatomical noise in contrast-enhanced digital mammography. Part I. Single-energy imaging. *Medical Physics,* 40, 51910.

Hill, M. L., Mainprize, J. G., Puong, S., Carton, A.-K., Iordache, R., Muller, S. & Yaffe, M. J. 2012. Impact of image acquisition timing on image quality for dual-energy contrast-enhanced breast tomosynthesis. *Proc. of SPIE,* 8313, 831308-831309.

Houssami, N. & Hayes, D. 2009. Review of preoperative magnetic resonance imaging (MRI) in breast cancer: should MRI be performed on all women with newly diagnosed, early stage breast cancer? *CA: a cancer journal for clinicians,* 59, 290-302.

Hsieh, J. 2009. *Computed tomography: principles, design, artifacts, and recent advances*, SPIE.

Hu, Y.-H., Zhao, B. & Zhao, W. 2008. Image artifacts in digital breast tomosynthesis: investigation of the effects of system geometry and reconstruction parameters using a linear system approach. *Medical Physics,* 35, 5242-5252.

Hu, Y.-H. & Zhao, W. 2011. The Effect of Angular Dose Distribution on the Detection of Microcalcifications In Digital Breast Tomosynthesis. *Med. Phys.,* 38, 2455-2466.

Hu, Y., Scaduto, D. & Zhao, W. 2013. Optimization of clinical protocols for contrast enhanced breast imaging. *SPIE Medical Imaging,* 8668, 86680H-1-9.

Hu, Y. & Zhao, W. 2012. Experimental quantification of lesion detectability in contrast enhanced dual energy digital breast tomosynthesis. *SPIE Medical Imaging,* 8313.

Hudson, H. M. & Larkin, R. S. 1994. Accelerated Image-Reconstruction Using Ordered Subsets of Projection Data. *IEEE Transactions on Medical Imaging,* 13, 601-609.

Hutton, B., Buvat, I. & Beekman, F. 2011. Review and current status of SPECT scatter correction. *Physics in Medicine and Biology*.

Icru 1989. Tissue Substitutes in Radiation Dosimetry and Measurement, Report 44 of the International Commission on Radiation Units and Measurements. Bethesda, MD.

Inscoe, C. R., Tucker, A. W., Zhou, O. Z. & Lu, J. 2013. Demonstration of a Scatter Correction Technique in Digital Breast Tomosynthesis *SPIE Medical Imaging,* 8668, .

Jong, R. A., Yaffe, M. J., Skarpathiotakis, M., Shumak, R. S., Danjoux, N. M., Gunesekara, A. & Plewes, D. B. 2003. Contrast-enhanced Digital Mammography: Initial Clinical Experience. *Radiology,* 228, 842-850.

Kamphuis, C., Beekman, F. & Hutton, B. 1999. Opimal Collimator Hot Dimensions for Half Cone-Beam Brain SPECT *Proc. of Int. Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine Meeting.*

Kay, S. 1998. *Fundamentals of Statistical Signal Processing: Detection Theory*.

Khurd, P. & Gindi, G. 2005a. Decision Strategies that Maximize the Area Under the LROC Curve. *IEEE Trans. Med. Imaging,* 24, 1626-1636.

Khurd, P. & Gindi, G. 2005b. Rapid Computation of LROC Figures of Merit Using Numerical Observers (for SPECT/PET Reconstruction). *IEEE Trans. Nucl. Sci.,* 52, 518-526.

Kolb, T., Lichy, J. & Newhouse, J. 2002. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. *Radiology,* 225, 165-175.

Krenning, E. P., Kwekkeboon, D. J., Oei, H. Y. & Al., E. 1994. Somatostatin-Receptor Scintigraphy in Gastroenteropancreatic Tumors: An Overview of European Results. *Ann NY Acad Sci,* 733, 416-424.

Kulkarni, S., Khurd, P., Zhou, L. & Gindi, G. 2009. Statistical Properties of SPECT MAP Reconstruction Incorporating Window Based Scatter Correction. *IEEE Transactions on Nuclear Science,* 56, 97-107.

Lamberts, S. W. J., Chayvialle, J. A. & Krenning, E. P. 1992. The Visualization of Gastroenteropancreatic Neuroendocrine Tumors. *Metabolism, Suppl,* 41, 111-115.

Lange, K. 1990. Convergence of EM image reconstruction algorithms with Gibbs smoothing. *IEEE Transactions on Medical Imaging,* 9, 439-446.

Lau, A. B. 2012. *Modeling Digital Breast Tomosynthesis Imaging Systems for Optimization Studies.* Ph.D., University of Chicago.

Lazos, D. & Williamson, J. F. 2012. Impact of Flat Panel-Imager Veiling Glare On Scatter-Estimation Accuracy And Image Quality of A commercial On-Board Cone-Beam CT Imaging System. *Med. Phys.,* 39, 5639-51.

Lee, C. J., Kupinski, M. A. & Volokh, L. 2013. Assessment of Cardiac Single-Photon Emission Computed Tomography Performance Using a Scanning Linear Observer. *Med. Phys.,* 40, 1-11.

Lewin, J., Hendrick, R., D'orsi, C., Isaacs, P., Moss, L., Karellas, A., Sisney, G., Kuni, C. & Cutter, G. 2001. Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. *Radiology,* 218, 873-880.

Lewin, J., Isaacs, P., Vance, V. & Larke, F. 2003. Dual-Energy Contrast-enhanced Digital Subtraction Mammography: Feasibility1. *Radiology,* 229, 261-68.

Lindfors, K., Boone, J., Nelson, T., Yang, K., Kwan, A. & Miller, D. 2008. Dedicated breast CT: initial clinical experience. *Radiology,* 246, 725-733.

Liu, Y. 2014. *Image reconstruction theory and implementation for low-dose X-ray computed tomography.* PhD, Stony Brook University.

Ljungberg, M. & Strand, S. 1989. A Monte Carlo Program for the Simulation of Scintillation Camera Characteristics. *Comput. Meth. Prog. Biomed.,* 29, 257-272.

Lu, Y. 2011. *Collimator Optimization for Single Photon Emission Computed Tomography Using Detection and Localization Tasks* M.S., State University of New York at Stony Brook.

Lu, Y., Chan, H.-P., Wei, J. & Hadjiiski, L. M. 2010. Selective-Diffusion Regularization for Enhancement of Microcalcifications in Digital Breast Tomosynthesis Reconstruction. *Med. Phys.,* 37, 6003.

Lu, Y., Lau, B., Hu, Y., Zhao, W. & Gindi, G. 2014a. A simple scatter correction method for dual energy contrast-enhanced digital breast tomosynthesis. *SPIE Medical Imaging,* 9033, .

Lu, Y., Peng, B., Scaduto, A. D., Zhao, W. & Gindi, G. 2014b. Application of the ordered-subsets transmission reconstruction algorithm to contrast-enhanced dual-energy digital breast tomosynthesis. *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference IEEE NSS/MIC*.

Lu, Y., Zhang, H., Liang, Z. R. & Gindi, G. 2013. An Approach to System Optimization for X-Ray Photon-Counting Systems Using Performance on a Detection/Localization Task. *2013 Ieee Nuclear Science Symposium and Medical Imaging Conference (Nss/Mic)*.

Lu, Y. H., Chen, L. & Gindi, G. 2014c. Collimator performance evaluation for In-111 SPECT using a detection/localization task. *Physics in Medicine and Biology*, 59, 679-696.

Mähler, E., Sundström, T., Axelsson, J. & Larsson, A. 2012. Detecting Small Liver Tumors With In-111 Pentetreotide SPECT—A Collimator Study Based on Monte Carlo Simulations. *IEEE Trans. Nucl. Sci.*, 59, 47-53.

Mariani, G., Bruselli, L., Kuwert, T., Kim, E. E., Flotats, A., Israel, O., Dondi, M. & Watanabe, N. 2010. A review on the clinical uses of SPECT/CT. *Eur J Nucl Med Mol Imaging*, 37, 1959-85.

Mertelmeier, T., Orman, J., Haerer, W. & Dudam, M. K. 2006. Optimizing filtered backprojection reconstruction for a breast tomosynthesis prototype device. *Proc. SPIE*, 6142, 61420F.

Metz, C. 1978. Basic Principles of ROC Analysis. *Seminars in Nuclear Medicine*, 8, 293-298.

Moore, S. C., Kijewski, M. F. & Fakhri, G. E. 2005. Collimator Optimization for Detection and Quantitation Tasks: Application to Gallium-67 Imaging. *IEEE Trans. Med. Imaging*, 24, 1347-1356.

Moore, S. C., Vries, D. J., Penney, B. C., Mueller, S. P. & Kijewski, M. F. 1998. Design of A Collimator for Imaging In-111. *In:* LJUNGBERG, M., STRAND, S. & KING, M. A. (eds.) *Monte Carlo Calculations in Nuclear Medicine Applications in Diagnostic Imaging*. UK: Institute of Physics Publishing

Niu, T. & Zhu, L. 2011. Scatter Correction for Full-fan Volumetric CT Using A Stationary Beam Blocker in A Single Full Scan. *Med. Phys.*, 38, 6027-38.

Ogawa, K., Harata, Y., Ichihara, T., Kubo, A. & Hashimoto, S. 1991. A Practical Method for Position-Dependent Compton-Scatter Correction in Single Photon Emission CT. *IEEE Trans. Med. Imaging*, 10, 408-412.

Paik, D., Beaulieu, C., Rubin, G., Acar, B., Jeffrey, R., Yee, J., Dey, J. & Napel, S. 2004. Surface Normal Overlap: A Computer-Aided Detection Algorithm with Application to Colonic Polyps and Lung Nodules in Helical CT. *IEEE Trans. Med. Imaging,* 23, 661-675.

Park, S., Jennings, R., Liu, H., Badano, A. & Myers, K. J. 2009. A Statistical, Task-based Evaluation Method for Three-Dimensional X-Ray Breast Imaging Systems Using Variable-Background Phantoms *Med. Phys.*, 37.

Parker & Anthony, J. 2011. *Joint Program in Nuclear Medicine* [Online]. Available: http://www.med.harvard.edu/JPNM/CH/Sept2303/BoneSPECT.jpg.

Paul, N. S., Siewerdsen, J. H., Patsicis, D. & Chung, T. B. 2007. Investigating the low-dose limits of multidetector CT in lung nodule surveillance. *Medical Physics,* 34, 3587-3595.

Piyavisetpat, N., Aquino, S. L., Hahn, P. E., Halpern, E. F. & Thrall, J. H. 2005. Small incidental pulmonary nodules - How useful is short-term interval CT follow-up? *Journal of Thoracic Imaging,* 20, 5-9.

Pratx, G. & Xing, L. 2011. GPU computing in medical physics: a review. *Medical Physics,* 38, 2685-2697.

Puong, S., Bouchevreau, X., Duchateau, N., Iordache, R. & Muller, S. 2008. Optimization of beam parameters and iodine quantification in dual-energy contrast enhanced digital breast tomosynthesis. *SPIE Medical Imaging,* 6913.

Puong, S., Patoureaux, F., Iordache, R., Bouchevreau, X. & Muller, S. 2007. Dual-energy contrast enhanced digital breast tomosynthesis: concept, method, and evaluation on phantoms. *SPIE Medical Imaging,* 6510, .

Ren, B., Ruth, C., Zhang, Y., Smith, A., Kennedy, D., O'keefe, B., Shaw, I., Williams, C., Ye, Z., Ingal, E., Polischuk, B. & Jing, Z. 2013. Dual energy iodine contrast imaging with mammography and tomosynthesis. *SPIE,* 8668.

Salvagnini, E., Bosmans, H., Struelens, L. & Marshall, N. 2012. Quantification of scattered radiation in projection mammography: four practical methods compared. *Medical Physics,* 39, 3167-3180.

Samei, E. & Saunders, R. 2011. Dual-energy contrast-enhanced breast tomosynthesis: optimization of beam quality for dose and image quality. *Physics in Medicine and Biology,* 56, 6359-6378.

Schlomka, J., Roessl, E., Dorscheid, R., Dill, S., Martens, G., Istel, T., Bäumer, C., Herrmann, C., Steadman, R., Zeitler, G., Livne, A. & Proksa, R. 2008. Experimental Feasibility of Multi-Energy Photon-Counting K-edge Imaging in Pre-Clinical Computed Tomography. *Phys. Med. Biol.*, 53, 4031.

Sechopoulos, I. 2013. A review of breast tomosynthesis. Part I. The image acquisition process. *Medical Physics,* 40, 14301.

Sechopoulos, I., Suryanarayanan, S., Vedantham, S., D'orsi, C. & Karellas, A. 2007a. Computation of the glandular radiation dose in digital tomosynthesis of the breast. *Med Phys,* 34, 221-32.

Sechopoulos, I., Suryanarayanan, S., Vedantham, S., J. D'orsi, C. & Karellas, A. 2007b. Scatter radiation in digital tomosynthesis of the breast. *Medical Physics,* 34, 564-76.

Segui, J. & Zhao, W. 2006. Amorphous selenium flat panel detectors for digital mammography: validation of a NPWE model observer with CDMAM observer performance experiments. *Medical Physics,* 33, 3711-3722.

Shieh, C. C., Kipritidis, J., O'brien, R. T., Cooper, B. J., Kuncic, Z. & Keall, P. J. 2015. Improving thoracic four-dimensional cone-beam CT reconstruction with anatomical-adaptive image regularization (AAIR). *Physics in Medicine and Biology,* 60, 841-868.

Siddon, R. L. 1984. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med. Phys.*, 12, 252-255.

Sidky, E. Y., Pan, X., Reiser, I. S. & Nishikawa, R. M. 2009. Enhanced imaging of microcalcifications in digital Breast tomosynthesis through improved image-reconstruction algorithms. *Med. Phys.*, 33, 4920-4932.

Siewerdsen, J. H., Daly, M. J., Bakhtiar, B., Moseley, D. J., Richard, S., Keller, H. & Jaffray, D. A. 2006. A simple, direct method for x-ray scatter estimation and correction in digital radiography and cone-beam CT. *Med. Phys.*, 33, 187-97.

Silverman, J. D., Paul, N. S. & Siewerdsen, J. H. 2009. Investigation of lung nodule detectability in low-dose 320-slice computed tomography. *Medical Physics,* 36, 1700-1710.

Sisniega, A., Zbijewski, W., Xu, J., Dang, H., Stayman, J. W., Yorkston, J., Aygun, N., Koliatsos, V. & Siewerdsen, J. H. 2015. High-Fidelity Artifact Correction for Cone-beam CT Imaging of the Brain. *Phys. Med. Biol.*, 60, 1415-39.

Skarpathiotakis, M., Yaffe, M. J., Bloomquist, A. K., Rico, D., Muller, S., Rick, A. & Jeunehomme, F. 2002. Development of contrast digital mammography. *Med. Phys.,* 29, 2419-26.

Smith-Bindman, R., Ballard-Barbash, R., Miglioretti, D., Patnick, J. & Kerlikowske, K. 2005. Comparing the performance of mammography screening in the USA and the UK. *Journal of medical screening,* 12, 50-54.

Sone, S., Takashima, S., Li, F., Yang, Z. G., Honda, T., Maruyama, Y., Hasegawa, M., Yamanda, T., Kubo, K., Hanamura, K. & Asakura, K. 1998. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet,* 351, 1242-1245.

Sundin, A., Garske, U. & Orlefors, H. 2007. Nuclear Imaging of Neuroendocrine Tumours. *Best Pract. Res. Clin. Endocrinol. Metab.,* 21, 69-85.

Swensson, R. 1996. Unified Measurement of Observer Performance in Detecting and Localizing Target Objects on Images. *Medical Physics,* 23, 1709-1725.

Tian, Z., Jia, X., Yuan, K. H., Pan, T. S. & Jiang, S. B. 2011. Low-dose CT reconstruction via edge-preserving total variation regularization. *Physics in Medicine and Biology,* 56, 5949-5967.

Van Trees, H., Bell, K. L. & Tian, Z. 2013. *Detection Estimation and Modulation Theory, 2nd Edition, Part I, Detection, Estimation, and Filtering Theory,* Wiley.

Wang, G. & Jiang, M. 2004. Ordered-subset simultaneous algebraic reconstruction techniques (OS-SART). *J. of X-Ray Science and Technology* 12, 169-177.

Wang, W. 1997. *Noise analysis of Bayesian reconstruction algorithms in emission tomography.* Ph.D., State University of New York at Stony Brook.

Wernick, M. N. & Aarsvold, J. 2004. *Emission Tomography: the Fundamentals of PET and SPECT,* Elsevier Academic Press.

Westcott, J., Barnden, L. & Lichtenstein, M. 2007. High Sensitivity versus General Purpose Collimator SPECT Imaging Using 3D Reconstruction Algorithm. *37th Ann. Scientific Meeting of the Australian and New Zealand Society of Nuclear Medicine.*

Wollenweber, S. D., Tsui, B. M. W., Lalush, D. S., Frey, E. C., Lacroix, K. J. & Gullberg, G. 1999. Comparison of Hotelling Observer Models and Human Observers in Detect Detection from Myocardial SPECT Imaging. *IEEE Trans. Nucl. Sci.,* 46, 2098-103.

Wu, G., Mainprize, J. G., Boone, J. M. & Yaffe, M. J. 2009. Evaluation of Scatter Effects on Image Quality for Breast Tomosynthesis. *Med. Phys.*, 36, 4425-32.

Wu, T., Moore, R. H., Rafferty, E. A. & Kopans, D. B. 2004. A Comparison of Reconstruction Algorithms for Breast Tomosynthesis. *Med. Phys.*, 31, 2636-47.

Yale University School of Medicine. 2004. *Cardiac Anatomy as Seen in SPECT Nuclear Imaging* [Online]. Available: http://www.yale.edu/imaging/techniques/spect_anatomy/index.html.

Yang, K., Burkett Jr, G. & Boone, J. M. 2014. A Breast-Specific, Negligible-Dose Scatter Correction Technique for Dedicated Cone-Beam Breast CT: A Physics-Based Approach to Improve Hounsfield Unit Accuracy. *Phys. Med. Biol.*, 59, 6487-6505.

Yankaskas, B. C., Cleveland, R. J., Schell, M. J. & Kozar, R. 2001. Association of recall rates with sensitivity and positive predictive values of screening mammography. *AJR Am J Roentgenol,* 177, 543-9.

Yankelevitz, D. F., Gupta, R., Zhao, B. & Henschke, C. I. 1999. Small pulmonary nodules: Evaluation with repeat CT - Preliminary experience. *Radiology,* 212, 561-566.

Zeng, G. & Gullberg, G. 2002. A Channelized-Hotelling-Trace Collimator Design Method Based on Reconstruction Rather than Projections *IEEE Trans. Nucl. Sci.*, 49, 2155-2158.

Zhang, Y., Chan, H. P., Sahiner, B., Wei, J., Goodsitt, M. M., Hadjiiski, L. M., Ge, J. & Zhou, C. 2006. A comparative study of limited-angle cone-beam reconstruction methods for breast tomosynthesis. *Med Phys*, 33, 3781-95.

Zhao, B. 2007. *Breast tomosynthesis with amorphous selenium digital flat panel detector.* Ph.D., Stony Brook University.

Zhao, B. & Zhao, W. 2005. Temporal performance of amorphous selenium mammography detectors. *Medical Physics,* 32, 128-136.

Zhao, B. & Zhao, W. 2008. Imaging performance of an amorphous selenium digital mammography detector in a breast tomosynthesis system. *Medical Physics,* 35, 1978-1987.

Zhao, W., Brunner, S., Niu, K., Schafer, S., Royalty, K. & Chen, G.-H. 2015. Patient-Specific Scatter Correction for Flat-Panel Detector-Based Cone-Beam CT Imaging. *Phys. Med. Biol.,* 60, 1339-65.

Zhou, L. 2007. *Low-Contrast Lesion Detection in Tomosynthetic Breast Imaging.* Ph.D. , Stony Brook University.

Zhou, L. & Gindi, G. 2009. Collimator Optimization in SPECT Based on a Joint Detection and Localization Task. *Physics in Medicine and Biology,* 54, 4423-4437.

Zhou, L., Khurd, P., Kulkarni, S., Rangarajan, A. & Gindi, G. 2008. Aperture Optimization in Emission Imaging Using Ideal Observers for Joint Detection and Localization. *Physics in Medicine and Biology,* 53, 2019-2034.