TASK-BASED IMAGE QUALITY ASSESSMENT IN X-RAY COMPUTED TOMOGRAPHY

by

Hsin-Wu Tseng

A Dissertation Submitted to the Faculty of the

COLLEGE OF OPTICAL SCIENCES

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2015

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Hsin-Wu Tseng
entitled Task-Based Image Quality Assessment In X-Ray Computed Tomography
and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Date: 16 November 2015
Matthew A. Kupinski

Date: 16 November 2015
Jiahua Fan

Date: 16 November 2015
Eric W. Clarkson

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.
I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: 16 November 2015
Dissertation Director: Matthew A. Kupinski

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Hsin-Wu Tseng

ACKNOWLEDGEMENTS

Behind the completion of my Ph.D, there is a team that has supported me since the first day I came to the United States of America. I thank God to bring them into my life while I pursued my Ph.D. I would like to take this opportunity to acknowledge these people. I would like to appreciate my older brother, Prof. Yuan-Chieh Tseng, who encouraged me to pursue my Ph.D when I hesitated about this dream. I am grateful for every educational support from my parents as well as their every pray every morning and every night. I also appreciated the encouragements from Tseng families when I have difficult time to find a research group and lost my dearest grandmom. I thank my wife, Dr. Jieyan Zhu, who offered her time unselfishly to company with me in the time of my preliminary exam and my four times summer internship in GE Healthcare (Waukesha, WI). I also appreciate the support from Zhu families who make me feel not lonely in my Ph.D life. I also would like to sincerely thank every person who prayed for me when I had troubles in my research.

This research would not be possible without the academic knowledge and guidance provided by my advisor, Prof. Matthew Kupinski, and co-advisor, Dr. Jiahua Fan from GE Healthcare. Thank you Prof. Kupinski for giving me this chance to join your group and providing me the opportunity of joining the joint projects with GE. Thank you for illustrating the concepts of model observers and the important simulation skills to me. Thank you for teaching me how to make plans feasible. I also appreciate your assistance on every publication. Thank you Dr. Fan for creating opportunities of summer internship every year and putting many efforts on our joint projects. These internship experiences and joint projects not only open a window for me to meet several experts in different CT fields but also provided me a unique chance to operate CT scanners (HD and Revolution). Thank you for teaching me how to design practical experiments with a discreet attitude. Many thanks to every member of the advanced scientific and engineering team of GE Healthcare for help on my experiments. Thank you Prof. Eric Clarkson for serving on my dissertation committee. I appreciate your time of explaining some ideas of noise models in your class. Additionally, I appreciate Prof. Harrison Barrett for his dedications to imaging sciences and being a good example of a scientist for every student in CGRI.

Finally, the most significant acknowledgement of all goes to my grandparents. Every descendant benefits from you because you used your whole life to provide a complete house and family. Thank you for raising me with so many efforts. Without your sacrifices, there is no chance for me to start my dream. During this Ph.D life, sometimes I really wanted to give up. The only motivation that sustained me to finish this dream was that I want to win honor for you.

DEDICATION

To my grandparents, my parents, my brother,
my wife, Tseng families, Zhu families,
and those who pray for me.

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF FIGURES – *Continued*

LIST OF TABLES

ABSTRACT

In X-Ray CT, there is always a desire to maintain the image quality while reducing the radiation dose. Recently several dose reduction approaches in both software and hardware have been developed to achieve the goal of making radiation as low as possible.Thus, the assessment of image quality becomes an important factor for routine quality control of medical X-Ray devices. In this work, task-based image quality measurements using model observers were used to evaluate the performance of X-Ray CT systems. To evaluate the dose reduction ability, detection tasks as well as combined detection and estimation tasks were considered. In detection tasks and combined detection and estimation tasks, the channelized Hotelling observer (CHO) and channelized scanning linear observer (CSLO) (with Dense Difference of Gauss channels) were employed respectively. They were used to evaluate the dose reduction capability of the iterative reconstruction algorithm developed by GE compared to the traditional reconstruction algorithm, filtered backprojection (FBP). Additionally, CHO and CSLO were also used for optimization of CT protocols. Our methods were also applied to Cardiac CT systems for temporal resolution evaluations. Two reconstruction algorithms, FBP and the motion correction algorithm, Snapshot Freeze (SSF), operated at two heart-beating rates with two reconstruction windows were quantitatively evaluated using task-based measurements. Finally, due to the huge demand of data acquisitions in the conventional channelized model observers, a proposed High-Dose-Signal-LOOL CHO/CSLO (HL-CHO/CSLO) that could efficiently reduce the data requirement has also been investigated in the pure detection, and combined detection and estimation task. In all studies, the practicality and the use of real data is emphasized. The results of all these studies demonstrate the usefulness of the task-based measurements of image quality in X-Ray CT imaging.

CHAPTER 1

Introduction

1.1 General Introduction

There were more than 67 million computed tomography (CT) examinations performed each year in the United States (Mettler et al. (2009)), which is a remarkable increase compared to 3 million performed in the early 1980s. This extraordinary upsurge of using CT can be attributed to its improved diagnostic accuracy, ease of use, wide accessibility, and speed of data acquisition with modern CT systems. While the image quality in CT has improved substantially over the years, CT imaging is responsible for more than two thirds of total radiation doses in medical imaging even though CT accounts only for 11% of X-Ray examinations. Research has shown that some incidental cancer cases may be associated with CT scans. Patients benefit from CT technology, but its contribution to collective dose has significantly increased and a significant effort is required to control this trend and ensure that the benefits from the use of this technology outweigh the risks. On the other hand, even though the risks for an individual are small, the rapid increase of CT utilization has created some significant concern over the patient radiation dose.

In the past decade, several techniques for reducing the CT radiation dose have been developed. The challenge of reducing the dose is to maintain the image quality while the noise is increased at decreasing exposure level. Maintaining clinically acceptable image quality at a lower dose is the goal of many techniques currently employed for reducing the radiation dose. For example, X-Ray tube current modulation (McCollough, Bruesewitz, and Kofler (2006); Kalender, Wolf, and Suess (1999); Jakobs and Becker (2002)) is one such method used to reduce the dose. However, dose reduction is limited because the conventional filtered backprojection (FBP) reconstruction algorithms do not work well when electronic noise in the data acquisition system becomes a significant contributor to the measured signal (Silva, Lawder, Hara, Kujak, and Pavlicek (2010); Kalra, Maher, Sahani, Blake, Hahn, Avinash, Toth, Halperm, and Saini (2003)).

The other important issue is the temporal resolution in Cardiac CT. Within the past several years, continual improvements in CT technology have made visualization of coronary anatomy and better detection of stenotic lesions possible (Leber et al. (2005); Johnson et al. (2006)). However, even with these improvements, the motion artifacts are still important issues in Cardiac CT images. Motion artifacts can affect the accurate description of important coronary structures, and they might also increase the variability of coronary calcium scores (Mao et al. (2001); Horiguchi et al. (2002); Detrano et al. (2005)). These effects reduce the diagnostic performance for the detection of stenotic coronary lesions (Ropers et al. (2003); Raff et al. (2005)) and make efforts at characterizing the composition of coronary plaques more complicated (Leber et al. (2004)). Undoubtedly, all of these issues cause physicians to often evaluate the image quality to determine whether there is useful and meaningful information.

An efficient approach to solve radiation issues is the advanced iterative reconstruction algorithm (Marin et al. (2010)). Advanced iterative reconstruction algorithms have been well defined recently to overcome the problem of image noise that occurs in CT images when using FBP while the dose is reduced. Some research (Silva et al. (2010); Marin et al. (2010); Hara et al. (2009)) also shows that they possess the property of producing high quality images at lower dose radiation. In addition to having the ability of maintaining image quality with lower radiation dose, iterative reconstruction algorithms also provide a solution for correcting motion artifacts. Another way to improve the image quality and reduce dose is using protocol design. Especially in a contrast enhancement CT scan, a good protocol not only reduces the radiation dose but also reduces the amount of the required contrast agent (Nakayama et al. (2005)). On the other hand, to improve the temporal resolution in Cardiac CT, motion compensation approaches (Iatrou et al. (2010)) have been demonstrated to have significant improvements on coronary reconstruction.

No matter the dose reduction strategy, quantitative evaluation of image quality is necessary. Any approach of reducing radiation dose or optimization of a protocol should be performed within a clinical framework. To be meaningful, image quality assessments should be related to actual clinical performance. Direct determination of clinical performance is quite difficult, expensive and time-consuming. The results in these kinds of studies really depend upon the patient sample strongly. Thus, it is important to design practical task-based image quality measurements and use patient simulating phantoms which can mimic important disease-related structures.

To evaluate the image quality, traditional object metrics such as contrast-to-noise ratio (CNR) in a region of interest (ROI) and noise power spectrum (NPS) were quite useful during the early development of CT technology. However, these approaches were limited in sensitivity and became invalid after the introduction of iterative reconstruction in CT. The nonlinearity of iterative reconstruction brings a challenge in image quality assessment. To establish a bridge between medical physicists and radiologists and to save time and cost, mathematical model observers currently are employed to evaluate the image quality in the CT community.

In addition to giving the basic concepts of the physics of X-Ray CT, image reconstruction algorithms, the unit of radiation dose, image presentation, dose metric, and artifacts, the majority of this dissertation will be focused on two parts: (1) design proper tasks and select model observers to quantitively evaluate the image quality or the dose reduction of different CT imaging systems and (2) make these image quality algorithms more practical.

1.2 Fundamentals of X-Ray Physics

X-Ray is one kind of electromagnetic wave. The wavelengths of X-Rays range from a few picometers to a few nanometers. The energy of an X-Ray photon is proportional to its frequency ν, and it can be expressed by the following equation:

$$E = h\nu = \frac{hc}{\lambda} \tag{1.1}$$

where h is Plank constant ($6.63 \times 10^{-34}\ J\ s$), c is the speed of light ($\approx 3 \times 10^{8}\ m\ s^{-1}$), and λ is the wavelength. Thus, we can easily see that the shorter the wavelength, the greater the energy. Since X-Ray photons are produced by striking a target metal with high-speed electrons (accelerated across an electrode), the maximum possible energy of an X-Ray photon is equal to the kinetic energy of the electron. Therefore, conventionally the unit of energy of X-Ray photons is usually expressed in eV (e is the charge of an electron and V is the voltage of electrode). Then, the equation (1.1) can be expressed as

$$E = \frac{1.24 \times 10^{3} eV \times nm}{\lambda}. \tag{1.2}$$

As mentioned above, X-Ray photons are generated when a target material is bombarded by high-speed electrons. The target material, filament current and accelerating voltage all have significant effects on the final output of an X-Ray generator. There are some choices for target materials (Mo, Rh, W). Different materials result in different photon generation efficiency. When an electron collides with a target material, different types of interactions happen. Most of the interactions come from the energy transfer between high-speed electron and electrons that are knocked out from atoms. However, this type of interaction does not contribute to the production of X-Rays and eventually is converted to heat. For a typical X-Ray tube, 99% of energy becomes heat.

With very low probability (about 0.5% of the time), an electron comes into proximity of a positively charged nucleus in the target. The electron is decelerated by the Coulomb interaction and its trajectory is changed as well. Then an X-Ray photon with the same energy as the lost kinetic energy of the electron is generated. This kind of radiation is called bremsstrahlung which means braking radiation in German.

The X-Ray energy produced during the bremsstrahlung varies with the distance between the incident electron and the target nucleus because the Coulombic force is inverse proportional to the square of the distance. The shorter the distance, the higher the energy of X-Ray photon. A direct collision of an bombarding electron with the nucleus is very rare because the cross section of the nucleus is very small.

An unfiltered bremsstrahlung spectrum (Fig. 1.1 (a)) depicts the distribution of X-Ray photons verse energy. The ramp-shaped distribution indicates that the number of lower-X-Ray energy photons is much greater than the number of higher-energy photons. The highest X-Ray energy is determined by the peak voltage (kVp) applied across the current tube (Fig. 1.2).The lower-energy part which is not useful to diagnosis is usually removed by a filter (Fig. 1.1 (b)).

In addition to the bremsstrahlung radiation and the radiation caused by direct collision, the other important radiation is the characteristic radiation. In the atom, each electron has a binding energy that depends on the shell where it resides. The closet shell is the K shell which has the highest binding energy. The next highest binding energy is L shell followed by the M shell. When an incident electron has an energy larger than the binding energy of an electron of the target atom, a collisional interaction is possible, causing the shell electron to be ejected, thus ionizing the atom. The unfilled shell is unstable, and an electron that resides on the outer shell will fill the vacancy. When the outer shell electron transitions to the lower unfilled shell, a characteristic X-Ray photon with the energy equal to the difference between the binding energies of the shells can be released. For different materials, binding energies are different and unique. Thus, the emitted X-Ray photons have discrete energies that are characteristic of the specific material. Among all the transitions, the most useful characteristic X-Rays in the diagnostic energy range are from other shells transmitted to K-shell. Table 1.1 lists common target materials and the associated binding energy of their K, L, and M shell electrons.

Some important concepts of the process of generating X-Rays and radiation spectrum have been described above. The interaction between X-Ray and matter and details of X-Ray tubes, cathode, and anode, etc. is out of scope of this dissertation and will not be addressed here.

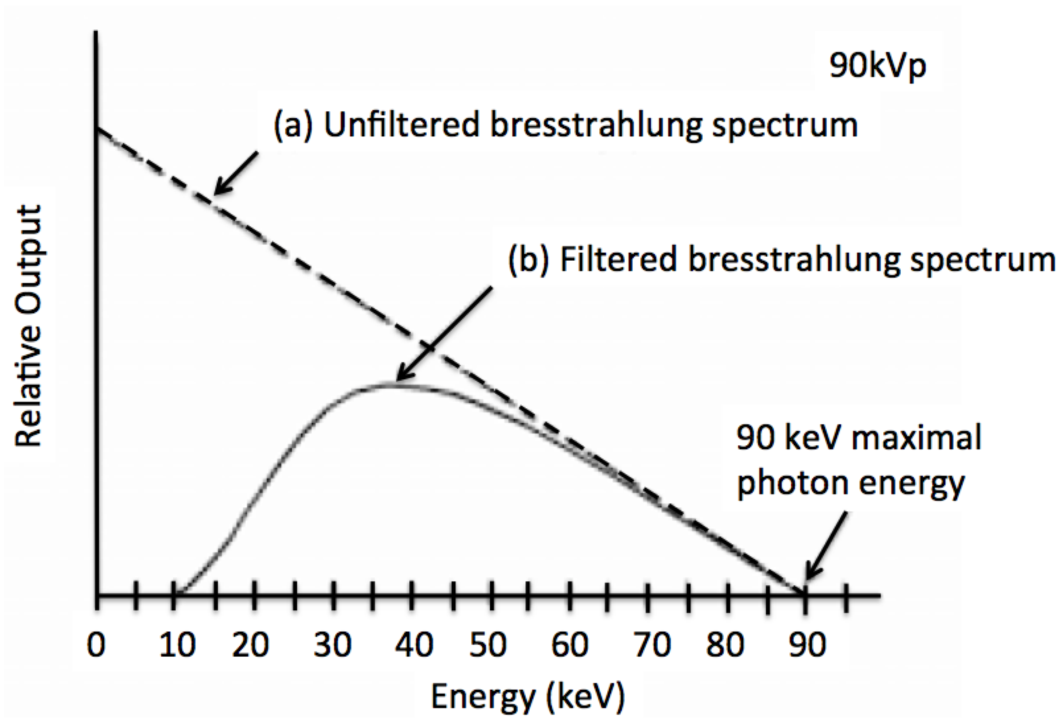

Figure 1.1: (a) Unfiltered bresstrahlung spectrum (b) Filtered bresstrahlung spectrum. (Courtesy of The Essential Physics of Medical Imaging.)

Table 1.1: Electron binding energies (keV) of common X-Ray tube target materials

Electron shell	Tungsten	Molybdenum	Rhodium
K	69.5	20.0	23.2
L	12.1/11.5/10.2	2.8/2.6/2.5	3.4/3.1/3.0
M	2.8-1.9	0.5-0.4	0.6-0.2

1.3 Computed Tomography

Computed tomography (CT) has been an invaluable diagnostic tool for clinical applications over the past decades. CT is the modality that made investigation of internal parts of the body possible. The first CT system was built in 1972, and CT technologies have matured greatly with improvements in computational technologies.

The mathematical principles of CT were first developed by Johann Radon in 1917. In Radon's work, he proved that an image of an unknown object could be produced from continuous projection data. The mathematical details will not be shown here because they are out of scope of this dissertation. Imagine that the three–dimensional (3D) anatomy of a patient is imaged with a plain film from two directions— posteroanterior and lateral projections. Like conventional radiography, the density at a given point on an image represents the attenuation properties within the patient along a line between the X-Ray focal point and the detector. The limitation here obviously is that the information associated with the dimension which is parallel to the beam is not obtained. If we acquire the information from these two directions, then the height, width, and depth of a lung on chest images can be defined. Therefore, complete thoracic information is obtainable if a series of 360 conventional radiographs acquired at 1-degree angular increments around the chest part of this patient. These 360 cross-sectional images, however, are not meant for human visualization.

To make acquired projection data useful, image reconstruction algorithms are necessary. After the reconstruction, CT images are produced and are usually presented as a series of slices along the long axis of the patient. The thickness of a CT image is usually thinner (0.625 mm to 5 mm) than a conventional X-Ray film. The two-dimensional (2D) array of pixels in the CT image is related to the 3D voxels. Voxels and pixels have the same dimension in plane, but voxels have slice-thickness dimension. Thus, each pixel in the CT image is actually the average properties of the X-Ray attenuation of the tissue/organ of the corresponding voxel.

The early available scan mode of CT was the step-and-shoot mode. However, the process of acceleration and deceleration of the step-and-shoot scan within a few millimeters easily caused the motion and deformation of internal organs of a patient. As a result, artifacts appear in the images. If multiple breath-holds are needed for scanning an organ, it is difficult to train patients to repeat the same breath-holding level to avoid the slice-to-slice misregistration. Another disadvantage of the step-and-shoot mode is that the long scan time of this mode results in some negative impacts for contrast enhancement examinations. This is due to the fact that the uptake and washout time of contrast agents are very short in bodies.

To overcome these difficulties in clinical studies, the helical or spiral scan mode was proposed (Nishimura and Miyazaki (1988); Kallender et al. (1989)). It is called a helical or spiral scan because the traveling path is like a helix if we consider the relative motion between a point on the gantry and a point on the patient. Projections are continuous when the patient is translated with a constant speed under a helical/spiral scan. Because of the faster speed and no acceleration and deceleration during the scan, this technique avoids artifacts and makes the utilization of contrast easier.

In addition to the improvement on the scan mode, multi-slice CT was also introduced in the last decades to satisfy an increased demand in the spatial resolution in the z direction (table direction). The main reason is the tradeoff between the slice thickness and the volume coverage for single-slice CT scanners. For instance, a typical z-direction coverage to cover the entire area from the celiac artery to the calves is between 90 and 120 cm, a length that is much longer than the width of a single-slice detector. Consider another example: to have optimal contrast enhancement and minimal respiratory motion from a patient, a thinner collimatior or a slow helical pitch is needed. However, either of these would cause other difficulties such as longer scan time, misregistration of breath-hold, unacceptable interruptions for the contrast enhancement studies, and low efficient utilization of the tube (thinner the collimation implies more photons blocked). These problems were solved when the z coverage technique was introduced into the CT systems. The z resolution makes the z-direction width of X-Ray beam independent of the slice thickness. This idea makes a wider X-Ray beam possible in thinner slices and also provides more efficient X-Ray tube utilization.

The CT scanners used in my studies were GE HD Discovery 750, and GE Optima CT 660 CT system. The Discovery CT 750 HD is a high-definition CT system. This CT scanner produces high definition images at up to 230 micron resolution. It also has been shown to produce 2.5 times more views per rotation which will improve the overall resolution. On the other hand, Optima CT 660 is one of the most energy efficient CT scanners available and has an “environmental design” that eases refurbishment and end-of-life recycling. The geometry and other features of scanners will not be addressed here. Most information on GE CT scanners is available on the GE website for interested readers. Both of them are now supported by an new developed iterative image reconstruction algorithm to achieve the low-dose imaging.

1.4 Cardiac Computed Tomography

Cardiac computed tomography (CT) is a new application of CT and has become in high demand during the last two decades. Improvements of technology such as faster rotation times, dual source, z-coverage, and smaller slice thickness fulfill the clinical studies of Cardiac CT and make Cardiac CT become a widespread tool. There are two major types of Cardiac CT applications: coronary artery calcification (CAC) and coronary artery imaging (CAI).

John Hunter(Oliver et al. (1964)) might be the first person who described the relationship between coronary artery calcification (CAC) and coronary death. The technology of X-Ray fluoroscopy was a method to detect calcium. However, the nature of fluoroscopy limited this technology because of the limited sensitivity. The usage of CT for calcification detection, in recent years, has obtained a lot of attention. The presence of calcium comes from a chronic process of injury and healing of the blood vessel walls. The amount of calcium shown in the arteries could be an important sign of coronary disease that might cause a heart attack. During the process of a CAC screen, CT images scanned at low dose selected at the diastole phase of the heart are reconstructed to minimize the cardiac motion artifacts. This can be accomplished by electrocardiogram (ECG)-based data acquisition (see following). According to these reconstructed images, calcium scores are obtained for a quantitative estimate of the stenosis. A successful CAC scoring relies upon its sensitivity and reproducibility.

The second type of application of cardiac imaging in CT is coronary artery imaging (CAI).This type of application typically needs the help of contrast agents because the purpose of CAI is to give aid to physicians to visualize the structures of the heart and to detect plaque and any other abnormalities. To have a good visualization in a small vessel, there are at least two challenges for CT systems. First, the system has to freeze the heart motion. Secondly, the system must have high spatial resolution in order to estimate the size of the vessel. This implies that the high temporal and spatial resolution are both necessary.

To minimize the artifacts and degradation in images, ECG signals are usually used in the data acquisition of CAC and CAI. A heart-beating cycle has two phases where the cardiac motion is relative small: the end-systolic phase and the end-diastolic phases. On an ECG trace, the region between 70% to 75% of the R-R interval (the elapsing time between two consecutive R waves in the ECG) refers to the mid-diastolic phase, and the region between 30% to 35% of the R-R interval corresponds to the end-systolic phase. Many research results (Seifarth et al. (2007)) have suggested that data acquired during the diastolic phase could provide better image quality for the patients with slow heart rates, and the end-systole phase could give a better image quality for the patients with higher heart rates.

The details of the parameter setup to optimize the cardiac scan will not be addressed here. This dissertation will focus on the image quality evaluation in Cardiac CT. In chapter 5, a task-based image quality assessment method will be used for the temporal resolution evaluations. The coronary artery images reconstructed by two different reconstruction algorithms and selected at two different reconstruction windows for a low and a high heart rate will be evaluated.

1.5 Image Reconstructions in CT

The most commonly used image reconstruction in CT is filtered backprojection (FBP). From the name, one can easily understand that there are two steps in the reconstruction: filtering the projection data followed by backprojection. The FBP algorithm could be described as the following: (1) Find the 1D Fourier transform of the projection data at every view angle. (2) Multiply the 1D Fourier transform of the projection data at every view angle with a ramp filter. (3) Find the 1D inverse Fourier transform of the output of step (2) for every view angle. (4) Integrate the output of step (3) over angles to obtain the final reconstructed image.

The other method to implement FBP is to use convolution. The details of derivations of FBP and different implementations are discussed in references (Herman (1980); Kak and Slaney (1988)).

Although FBP is well known for its speed and robust image quality, there are several disadvantages in this algorithm. For example, small metallic objects may induce streaking artifacts in the FBP image. FBP is not easily adaptable to missing data and partial occlusion effects. FBP images suffer from noise and artifact contaminations especially in low radiation dose conditions.

On the other hand, iterative methods use statistical models of the noise to improve the image in each iteration step. Various techniques have been offered by major vendors. Typically, the goal is to find the relationship between the projection and image data and the prior distribution of images. The optimization is very difficult to solve analytically so it requires iterative technieques.

Conventionally, a CT iterative reconstruction (IR) requires accurate modeling of the system optics and geometry, noise, image. Excellent image quality has been achieved at the cost of computational intensity (Hsieh et al. (2013)). To achieve real time image reconstruction speeds while also maintaining the dose-saving capability, GE Healthcare (Waukesha, WI) has designed an approach that uses simplified models of the CT imaging systems and greatly reduces the computational cost. The model focuses mainly on the system noise statistics, imaging object, physics and their interactions. With low flux and noise compensation implemented in this algorithm, the image quality is expected to equal to that of FBP reconstructed images but at a fraction of the dose level. The detailed description of this algorithm will be presented in a separate publication document. The focus here is to determine the amount of the dose saving compared to FBP this algorithm can provide.

The other problem that FBP suffers from is motion artifacts. Motion (patient, cardiac) causes blurring and double images. Faster scanners could reduce motion artifacts because the patient/heart has less time to move/beat during the data acquisition. Another solution is to use algorithmic compensation: (1) iterative methods that utilize a subset of the data and some priori knowledge (Chen et al. (2010)) and (2) motion compensation approaches (Chun and Fessler (2009); Taguchi et al. (2007); Cho et al. (2012)). Again, this dissertation will focus on the comparison between FBP and motion compensation approaches developed in GE Healthcare, and more details can be found in the references.

1.6 CT Numbers and Hounsfield Units

The typical range of the X-Ray photons generated for medical CT is roughly between 20 keV to 140 keV. There are three major interactions in which X-Rays interact with matter in this range: the photoelectric effect, the Compton effect, and the coherent scattering. The net effect of these interactions is that some of the X-Ray photons are absorbed or scattered. In other words, X-Ray photons are attenuated when they pass through materials. The Beer-Lambert law is used to describe the relationship between the transmitted rays, incident rays, and attenuation coefficients of the material:

$$I = I_0 e^{-\mu l} \tag{1.3}$$

where I and I_0 is the intensity of the transmitted and incident rays, respectively, μ is the summation of the attenuation coefficients of photoelectric, Compton, and coherent interactions of the material, and l is thickness of the material. Note, μ can be a function of X-Ray energy.

The μ values for water and muscle are 0.1928 cm^{-1} and 0.1916 cm^{-1} at the average photon energy, 70 keV, when the tube is operated at 120 kVp. Obviously, the difference is very small. To properly display this information, the difference between different materials needs to be enhanced. The intensity scale, CT number, used in reconstructed CT image is defined as

$$CTnumber = \frac{\mu - \mu_{water}}{\mu_{water}} \times 1000 \tag{1.4}$$

where μ_{water} is the attenuation coefficient of water. The unit of CT number is the Hounsfield unit (HU). This normalization results in CT numbers ranging from -1000 to +3000, where -1000 refers to air, soft tissues have range from -300 to -100, water is 0, and dense bone and contrast agent corresponds to +3000.

1.7 Image Presentation

Due to the wide range of CT number, it is necessary to have a modification before displaying the image on a standard grayscale monitor or film. When a reconstruction image with CT number ranges from -1000 HU to 1700 HU, it will be difficult to see the intensity variation if this dynamic range is linearly mapped to the range of the display device (0 to 255). To make the displayed image more visually appealing, a CT image is usually displayed with a modified grayscale:

$$p(x,y)_{display} = \begin{cases} 0, & p(x,y) \leq L - \frac{W}{2} \\ \frac{p(x,y)-(L-\frac{W}{2})}{W} I_{max}, & L - \frac{W}{2} < p(x,y) \leq L + \frac{W}{2} \\ I_{max}, & p(x,y) > L + \frac{W}{2} \end{cases} \tag{1.5}$$

where $p(x,y)_{display}$ is the final displayed image, $p(x,y)$ is the original reconstructed image, W means the window width (WW), L refers to the window level (WL), and I_{max} is the maximum intensity scale of the display device (I_{max}=255 for 8-bit display device). The meaning of 1.5 is essentially to map the original intensity scale between $L - \frac{W}{2}$ and $L + \frac{W}{2}$ to the full scale of the display device. In this range ($L - \frac{W}{2}$ and $L + \frac{W}{2}$), the linear transformation is used. However, in most of cases, nonlinear mapping is more helpful to either enhance or de-emphasize the interesting features. The mapping function can in general be expressed as

$$p(x,y)_{nonlineardisplay} = \begin{cases} G[p(x,y)_{display}], & t_L \leq p(x,y)_{display} < t_H \\ p(x,y)_{display}, & otherwise \end{cases} \tag{1.6}$$

where $p(x,y)_{nonlineardisplay}$ is the displayed image after the nonlinear mapping, G is the nonlinear mapping function, and t_L and t_H are parameters of the nonlinear mapping function. In my studies, the WW and WL were selected at default values in the GE consoles for the data acquisition.

1.8 Dose Metric - CTDI volume

U.S. Food and Drug Administration (FDA) has published guidelines to provide some recommendations on how to optimize CT protocols (Feigal (2002)), especially for pediatric patients and small adults, to reduce radiation exposure. To have a concept of how much radiation was measured for a scan, one has to understand the methods of the dose measurement. Several dose descriptors have been used in the past. The most commonly used one is Computed Tomography Dose Index (CTDI). Different versions of CTDI have been used. CTDI was initially introduced by Shope et al (Shope et al. (1981)) in 1981. CTDI was used as a metric to quantify the radiation output from a CT examination consisting of multiple contiguous CT scans. The word “index” was specifically included in this name to distinguish the quantity from the radiation dose absorbed by a patient. This method was adopted by the Center for Devices in Radiological Health (CDRH) of the FDA. U.S federal regulations require the CT measurement to report CTDI values measured in phantoms that are at least 14 cm long with a diameter 16 cm for head and 32 cm for body. PMMA phantoms are the most commonly used phantoms for dosimetry.

There were some issues from the scattered radiation, divergence of the radiation beam, and limitations from the efficiency of the beam collimation that cause delivered radiation to not be fully contained within the scanning volume. To address these problems, the $CTDI_{100}$ was developed. It was designed for calculation of the index for 100 mm along the length of a pencil ionization chamber. The weighted CTDI ($CTDI_w$) was then developed to solve the limitations of position dependence within the scanning plane. $CTDI_w$ represents a dose index that includes a weighted average of contributions from the central and peripheral of the phantoms within in the scanning plane. The weight factor is one-third and two-third for central and peripheral respectively.

The $CTDI_w$ is defined only for the step-and-shoot scan. The definition of $CTDI_w$ does not take into account the X-Ray exposure for the helical or spiral scan. One parameter that describes the speed of the motion of the patient table is the helical pitch. The helical pitch is defined as the ratio of traveling distance in one gantry rotation over the nominal beam width. The $CTDI_{vol}$ then was introduced to take into account the pitch of the helical acquisition. The $CTDI_{vol}$ represents the average delivered radiation and is calculated as the $CTDI_w$ divided by pitch. The unit of this value is expressed in mGy and is now displayed on most of CT consoles. In clinical environments, the scan range in z direction might vary significantly. Thus, another term, Dose-Length Production (DLP), is defined as the $CTDI_{vol}$ multiplied by scanning length. In our study, since the scanning lengths were all the same, we used $CTDI_{vol}$ as our dose metric.

1.9 Artifacts in CT

Artifacts are common in clinical CT. They can seriously degrade the quality of CT images and even sometimes obscure the pathology. CT artifacts originate from a wide range of sources. They can be separated into at least four types of artifacts: (1) Physics-Type artifacts: These come from the physical processes in the acquisition of CT data. (2) Patient-Type artifacts. This type of artifact is caused by the patient movement, heart beating, or the presence of metallic materials in or on the patients. (3) Scanner-Type artifacts: Artifacts are produced by defects in the scanner function. (4) Helical and Multi-Slice-Technology Type artifacts: Artifacts generated from the image reconstruction process.

Most of these artifacts can be avoided by calibration, careful patient positioning, and parameters/protocols setup. However, some of the artifacts, such as Poisson noise from lower radiation dose level and the heart beating, are not easily overcome. Poisson noise comes from the statistical error in low photon counts. It can result in random thin bright and dark streaks appearing along the direction of greatest attenuation. High contrast objects such as bone might be still visible with low dose level, but low contrast soft tissue boundaries usually become invisible. On the other hand, the rapid motion of heart generates artifacts which look similar to diseases and can lead to diagnostic errors. Techniques, such as electrocardiographic gating (ECG), have been developed to produce images by using partial data from the least cardiac motion region of the cardiac cycle. However, as mentioned before, the traditional algorithm, FBP, suffers from low radiation dose and the cardiac motion even when using ECG. Besides, protocol optimization related to the patient size and age etc. is not trivial. To overcome these issues, advanced image reconstruction algorithms and the proper selection of protocol are needed. To study the performance of different algorithms or different protocols, image quality evaluations are necessary. In the following section, the concepts and limitations of the traditional image quality metrics as well as task-based image quality metrics will be illustrated.

1.10 Image Quality

CT is a 3D imaging technology in which approaches for image quality evaluation have to be selected carefully. In this section, the traditional physical metrics will be described first. The limitations of these kinds of metrics will also be illustrated. Task-based image quality assessments will then be described in the second subsection.

1.10.1 Traditional Metrics

The traditional objective metrics are also called physical measurements. They are directly related to the physical properties of signal and noise of the imaging systems.

Some of them were developed in the spatial domain and others were designed to be used in the spatial–frequency domain because some features produce responses that are correlated spatially, whereas others produce responses independent of location.

1.10.1.1 Spatial Domain Metrics

A common image quality metric in spatial domain is contrast-to-noise ratio (CNR). CNR is defined as

$$CNR = \frac{\mu_{object} - \mu_{background}}{\sigma_{background}} \tag{1.7}$$

where μ_{object} and $\mu_{background}$ are the mean CT number of a object in the region of interest and the image background respectively and $\sigma_{background}$ is the background noise expressed as a standard deviation. CNR is not considered as a optimal image quality metric because the information of the size of the object and the pixel size of the image is not considered. Examples can be found in chapter 4.

In the following, we will see some useful image quality assessments in Fourier domain.

1.10.1.2 Spatial Frequency Domain Metrics

To distinguish two separate objects, the spatial resolution needs to be defined. The spatial resolution is directly related to the pixel size, the kernel used in the image reconstruction, and other hardware designed in the imaging systems. To understand image resolution, an image slice $I(x, y)$ and the input object $f(x, y)$ can be linked together by the point spread function. To simplify this derivation, the axial plane is considered. The relationship between $I(x, y)$ and $f(x, y)$ is expressed as

$$I(x, y) = \int\int f(x - x', y - y')PSF(x', y')dx'dy' \tag{1.8}$$

where PSF is the point spread function. The PSF is used for describing resolution properties of the imaging systems. The Fourier transform of the PSF is the optical transfer function (OTF) which is defined as

$$OTF(\mu, \nu) = \mathcal{F}\{PSF(x, y)\} \tag{1.9}$$

A more common way to estimate the spatial resolution is the modulation transfer function (MTF), and it is related to the OTF by the following equation

$$MTF(\mu,\nu) = \frac{|OTF(\mu,\nu)|}{|OTF(0,0)|} \tag{1.10}$$

The meaning of MTF is to measure how well frequencies are transferred through the system. It provides the information of responses of the system to different frequencies. The Fig. 1.3 shows two MTF curves. The flat one refers to the ideal system such that the system response is independent of the input frequency. Practically speaking, the MTF can be computed not only by the PSF, but also the line spread function (LSF) and the edge spread function (ESF) (Nikoloff and Wagner (1985))(Boone (2001))(Judy (1976)). More extensive details of practical implementations can be found in reference 23 (International Commission on Radiation Units and Meausrements. ICRU Report No.87 (2012)).

Similar to the resolution, image noise can be estimated in the Fourier domain. In a CT system, there are different sources of noise. Electronic noise comes from the detector readout. Quantum noise corresponds to the statistics of the limited quanta contributing to the image. The noise power spectrum (NPS) (or Weiner spectrum) is usually used as a image quality metric in the cases that the image statistics is stationary . It provides a description of the noise by giving the amplitude of the noise over the whole range of frequency in the image. To compute the NPS, an acquired homogenous region of interest (ROI) of images is required and the NPS is given by

$$NPS_{2D}(\mu,\nu) = \frac{\Delta x \Delta y}{L_x L_y}\frac{1}{N_{ROI}}\sum_{i=1}^{N}|\mathcal{F}\{ROI_i(x,y) - \overline{ROI_i}\}|^2 \tag{1.11}$$

where Δx, Δy are the pixel sizes in the x and y dimension, L_x, L_y are the two dimensions of the ROI, N is the number of ROIs used in the analysis, and $R\bar{O}I_i$ is the mean value of ith ROI. The relationship between σ and NPS can be expressed as

$$\sigma^2 = \int\int NPS_{2D}(\mu,\nu)d\mu d\nu \tag{1.12}$$

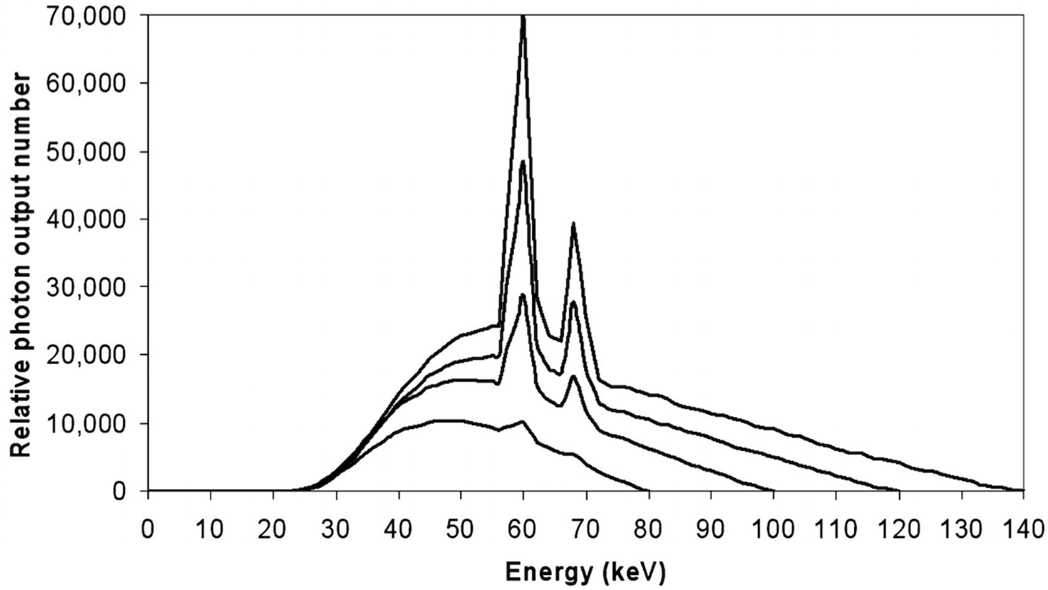

Figure 1.2: X-Ray spectrum: 80, 100, 120, and 140 kVp (Courtesy of the figure 5 of Seibert (2004)).

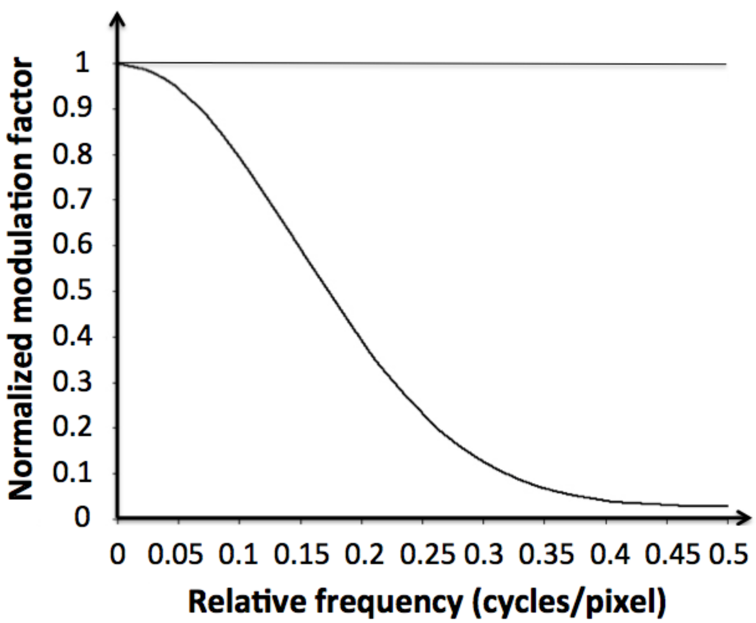

Figure 1.3: MTF. (Flat) Ideal system. (Curved) Actual system.

Although these metrics have been used in CT, they are not selected to be the image quality evaluation assessments in my study because the shift-invariance and linearity assumptions are not valid. The divergent X-Ray violates the shift-invariance and there is no linear property in iterative reconstruction algorithms. In the next section, task-based objective metrics will be illustrated. They are not limited to shift-invariance and linearity. Currently, they are being used for quantitative evaluation of image quality in CT systems.

1.10.2 Task-Based Objective Metrics

To evaluate image quality for different imaging systems, state-of-the-art medical image quality assessment methods that measure the ability to extract desired information from the images for clinically interesting tasks are now known to be good choices. A clinically interesting and relevant task might be a detection task that requires distinguishing normal cases from diseased cases. Another example of a relevant task is a estimation task that may provide information of the size and contrast of lesions. In other words, image quality in medical systems should be measured by an observer performing a task or tasks of clinical interest. In addition to the selection of the task, a description of the observer and the figure of merit are also key elements for task-based measurements. The observer is the person or algorithm that performs the task. Examples include human observers and model observers, which we will use in this work. Finally, the figure of merit is a way of measuring how well the observer performs the task. This figure of merit quantifies the ability of the observer to use the image data to perform the relevant task and is an objective measure of image quality.

1.10.2.1 Human Observer

A human observer is a very straightforward method for image quality evaluation. An example of a human observer for detection task is the two alternative forced choice (2-AFC). The observer has to choose one figure that has a signal–present from two options, even if he/she has to guess. The figure of merit here is the proportion of correct (PC) responses.

For 2-AFC, the PC is an estimate of the AUC value (the area under the receiver operating characteristic (ROC) curve) from human observer in ROC (Barrett et al. (1998)) paradigm. The relationship between detectability and PC can be expressed as

$$d' = \sqrt{2}\Phi(PC)^{-1} \tag{1.13}$$

where d' is the detectability. More details of the properties about d', relationship between d', and the signal to noise ration (SNR), are in the references related to these topics (Green and Swets (1966); Burgess (2011)).

There are some distinct disadvantages to using human observers such as time and cost, so mathematical model observers have become attractive alternatives. The basic ideas of model observers will be described in the following subsection.

1.10.2.2 Mathematical Model Observer

In general, there are two kinds of applications of model observers. One is in vision science, which people are more interested in the response of the human visual system (Geisler (2003, 2011)). The other is in medical imaging which will be illustrated in this section. This dissertation will be focused on the second application. Topics related to vision science are left for interested readers to explore.

Model observers are valuable tools in the objective assessment of image quality. Model observers allow for rapid computation of task-based figures of merit while avoiding complicated and expensive human psychophysical studies. Model observers in medical imaging are developed for two major purposes. The first one is to use a model observer to use either output images or projection images to obtain statistical information to help optimize the hardware system for a given task of interest. They are used to evaluate the image quality affected by tuning available parameters in systems in the presence of noise and other factors (Abbey and Boone (2008); He et al. (2008); Graff and Myers (2011); Lee et al. (2013)). Abbey and Boone used Markov-Chain Monte-Carlo (MCMC) based ideal observer (IO) devised by Kupinski et al. (Kupinski et al. (2003)) to evaluate an emerging breast CT device. He, Gaffo, and Frey applied the MCMC technique (Kupinski et al. (2003)) to estimate the ideal observer test statistic in the context of myocardial perfusion. Graff and Myers developed an ideal observer based methodology for assessing the magnetic resonance (MR) acquisition sequences. Lee, Kupinski, and Volokh evaluated the relationship between the image quality and the number of detectors using the scanning linear observer.

The second purpose is to evaluate and optimize software systems such as image reconstruction algorithms. For this purpose, people are more interested in the developing the model observers that mimic human behavior that can be used to predict human observer performance. The motivation of developing model observers that could predict human observer performance is that the variability in human observer performance is significant. It has been observed that different human observers can have very different performances. This varying performance might come from experience level. Another source that causes the variability is the "jitter" from the same human observer. Human observers might even provide inconsistent ratings on the same testing images.

All observers performing a binary classification task such as signal detection can be thought of as mapping the image data to a decision variable. To implement the model observers in real CT systems, the parameters of the underlying distributions are unknown so usually a number of samples are needed for "training" before "testing". However, the number of training images available is finite. The size of training sample sets, relative to the number of features, determines the accuracy of the statistical estimates. Quadratic and linear classification discriminant function both have been investigated (Fukunaga and Hayes (1989)). Fukunaga suggested that the linear classifier is more robust (less sensitive to parameter estimation errors) than the quadratic classifier. His work also theoretically and experimentally supported the results from Novak (Novak (1984)). The linear discriminant functions are relatively easy to compute and their performance is easy to summarize, and far less information regarding to data statistics is required. Thus, the work in this dissertation is focused on the linear observer models. The mapping, then, from image data to the discriminant function can be written as

$$\lambda(\boldsymbol{\theta}) = \boldsymbol{w}^t(\boldsymbol{\theta})\boldsymbol{g}, \tag{1.14}$$

where g is a reconstruction image, t is transpose operator, $\boldsymbol{\theta}$ is the parameter vectors such as the location, size and contrast of the signal, $\lambda(\boldsymbol{\theta})$ is the observers discriminant function, and the $\boldsymbol{w}(\boldsymbol{\theta})$ is the linear template. For the pure signal–detection task with signal location known, $\boldsymbol{\theta}$ is a constant. For signal detection and localization task, $\boldsymbol{\theta} = \boldsymbol{r}$ is varied and observer looks for a $\boldsymbol{r}_{loc}$ which gives maximum $\lambda(\boldsymbol{r})$ as shown in the following equation (Khurd and Gindi (2005))

$$\lambda_{max} = max_{\boldsymbol{r}}\lambda(\boldsymbol{r}) = \lambda(\boldsymbol{r}_{loc}) \tag{1.15}$$

To illustrate the concepts of model observers, we simplify the task and focus on the pure detection task i.e. λ is not function of signal properties. The optimal template, $w = w_{optlin}$, maximizes the SNR_λ (defined in equation 1.19) in the binary detection task when the data have equal covariance matrices under each hypothesis. The test statistic λ can be expressed as

$$\boldsymbol{w}_{optlin} = \boldsymbol{S}^{-1}_{\boldsymbol{g}_{trainer}}\boldsymbol{s} = \boldsymbol{S}^{-1}_{\boldsymbol{g}_{trainer}}\Delta\bar{\boldsymbol{g}}_{trainer} \tag{1.16}$$

$$\lambda = \boldsymbol{w}^t_{optlin}\boldsymbol{g}_{tester} \tag{1.17}$$

where $\boldsymbol{S}_{\boldsymbol{g}^{trainer}}$ is the sampling covariance matrix, $\boldsymbol{s}$ is the estimated signal, $\Delta\bar{\boldsymbol{g}}_{trainer}$ is the average difference in the data under two hypotheses and the subscript index, trainer and tester, refer to training data and testing data respectively. The demonstration of optimality of the above equation can be found in the reference (Barrett and Myers (2003)). $\boldsymbol{w}_{optlin}$ has been called the Hotelling discriminant. An observer who implements the optimal linear discriminant has been called the Hotelling observer (HO). Since the following paragraphs will be focused on the Hotelling observer, the subscript, opt lin, will be ignored for convenience.

For a pure detection task, two classes, signal–present and signal–absent, provide two test statistics. They are given by

$$\lambda_L = \begin{cases} \lambda_2 = \boldsymbol{w}^t \boldsymbol{g}^{tester,signalpresent}, & L = 2 \text{ signal present} \\ \lambda_1 = \boldsymbol{w}^t \boldsymbol{g}^{tester,signalabsent}, & L = 1 \text{ signal absent} \end{cases} \tag{1.18}$$

where subscript index, 1 and 2, means signal–absent and signal–present class respectively, and two distributions (Fig. 1.4) can be plotted. By varying the threshold and plotting true positive fraction verse false positive fraction (for example: A, B, C)(Barrett and Myers (2003)), a ROC curve (Fig. 1.5) can be generated. The AUC value can be used as a figure of merit (FOM) in the task. Another useful FOM is SNR_λ and it can be calculated by the means and variances of these two distributions:

$$SNR_\lambda = \frac{\bar{\lambda}_2 - \bar{\lambda}_1}{\sqrt{\frac{\sigma_2^2 + \sigma_1^1}{2}}} \tag{1.19}$$

where the subscript λ used here to distinguish this SNR from the conventional SNR related to the number of photon counted and the σ is the standard deviation of the test statistic. If the test statistic is normally distributed under the two hypotheses, SNR_λ is granted the special name d' and the the AUC can be easily calculated by

$$AUC = \frac{1}{2} + \frac{1}{2} erf(\frac{SNR_\lambda}{2}) \tag{1.20}$$

where erf means the error function.

The observer that is utilized by $\boldsymbol{S}_{\boldsymbol{g}_{trainer}}^{-1}$ is also called the prewhitening observer. The simple concept is described as follows. If the noise covariance matrix $\boldsymbol{S}_{\boldsymbol{g}_{trainer}}$ is not singular, we can define a new random variable $\boldsymbol{z}$ in terms of the square-root of the covariance matrix $\boldsymbol{S}_{\boldsymbol{g}_{trainer}}$:

$$\boldsymbol{z} = \boldsymbol{S}_{\boldsymbol{g}_{trainer}}^{-\frac{1}{2}} \boldsymbol{g}_{trainer} \tag{1.21}$$

such that

$$\Delta \bar{\boldsymbol{z}} = \boldsymbol{S}_{\boldsymbol{g}_{trainer}}^{-\frac{1}{2}} \Delta \bar{\boldsymbol{g}}_{trainer} \tag{1.22}$$

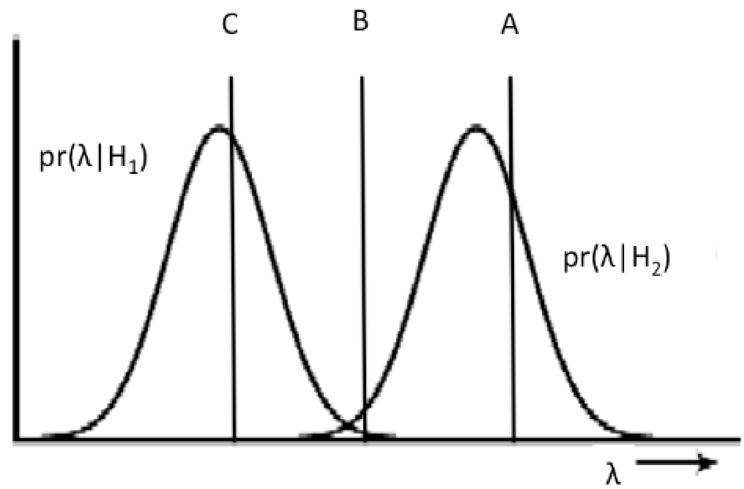

Figure 1.4: Distributions of test statistics and different thresholds.

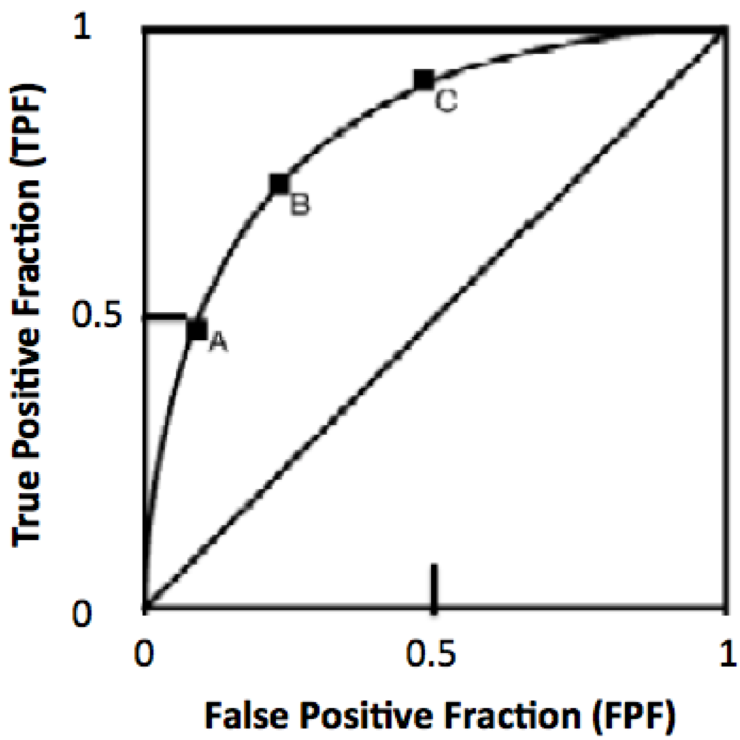

Figure 1.5: ROC curve with different thresholds.

and the test statistic becomes

$$\lambda = [\boldsymbol{S}^{-\frac{1}{2}}_{\boldsymbol{g}_{trainer}} \Delta \bar{\boldsymbol{g}}_{trainer}]^t [\boldsymbol{S}^{-\frac{1}{2}}_{\boldsymbol{g}_{trainer}} \boldsymbol{g}_{trainer}] = \Delta \bar{\boldsymbol{z}}^t \boldsymbol{z} \tag{1.23}$$

The covariance matrix of z can be calculated as

$$\boldsymbol{S}_{\boldsymbol{z}} =< [\boldsymbol{S}^{-\frac{1}{2}}_{\boldsymbol{g}_{trainer}} \Delta \bar{\boldsymbol{g}}_{trainer}][\boldsymbol{S}^{-\frac{1}{2}}_{\boldsymbol{g}_{trainer}} \Delta \bar{\boldsymbol{g}}_{trainer}]^t >= \boldsymbol{S}^{-\frac{1}{2}}_{\boldsymbol{g}_{trainer}} \boldsymbol{S}_{\boldsymbol{g}_{trainer}} \boldsymbol{S}^{-\frac{1}{2}}_{\boldsymbol{g}_{trainer}} = \boldsymbol{I} \tag{1.24}$$

where I is the identity matrix. The z, here, transformed by the training covariance matrix, is a uncorrelated random vector and this process is called prewhittening because the eigen-spectrum of the covariance matrix, $\boldsymbol{S}_{\boldsymbol{z}}$, is flat or white.

When the data are Gaussian, with equal covariance for each class, the HO is equal to the IO for tasks. However, when the data are not Gaussian, the ideal observer can be nonlinear, then full knowledge of the probability density functions of data is required for the likelihood ratio. This requirement is usually a hindrance to the calculation of the ideal decision variable (discriminant function), i.e. the likelihood ratio, in more realistic situations. On the other hand, the HO requires only the first- and second- order statistics of the data. Moreover, the HO has been found to be a good surrogate for the human observer (Rolland and Barrett (1992)). Because of these benefits, channelized Hotelling observers (CHO), were used in this dissertation for signal location known detection tasks and signal location unknown detection tasks. The reason for using channelization and the principles of the channelization will be described in the following section and chapter.

In addition to detection tasks, more complicated tasks such as a combination of detection, and estimation were studied as well. For the combined task, the channelized scanning linear observer (SLO) (Clarkson (2007)) was selected to evaluate image quality. The concept of scanning linear estimation is to calculate the mode of the posterior density or to maximize a posteriori (MAP) estimation. This requires a scan of parameter space to compare solutions and find the maximum. The general formula of MAP estimation is

$$\hat{\boldsymbol{\theta}}_{MAP} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \{pr(\theta|\boldsymbol{g})\} = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \left\{ \frac{pr(\boldsymbol{g}|\boldsymbol{\theta})pr(\boldsymbol{\theta})}{pr(\boldsymbol{g})} \right\} \tag{1.25}$$

where the $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$ are the parameters we are interested in and the estimated parameter respectively. $pr(\boldsymbol{g}|\boldsymbol{\theta})$ is the likelihood of data conditioned on parameters. To easily optimize the function, we will consider the Gaussian likelihood. Note this approximation does not imply that joint pdf $pr(\boldsymbol{g}, \boldsymbol{\theta})$ is also Gaussian. Based on this approximation, the conditioned likelihood $pr(\boldsymbol{g}|\boldsymbol{\theta})$ can be described by

$$pr(\boldsymbol{g}|\boldsymbol{\theta}) \cong \frac{1}{\sqrt{2\pi^M det(\boldsymbol{S}_{\boldsymbol{g}|\boldsymbol{\theta}})}} exp\Big[\frac{-1}{2}(\boldsymbol{g} - \bar{\boldsymbol{g}}(\boldsymbol{\theta}))^t \boldsymbol{S}_{\boldsymbol{g}|\boldsymbol{\theta}}^{-1}(\boldsymbol{g} - \bar{\boldsymbol{g}}(\boldsymbol{\theta}))\Big] \tag{1.26}$$

where $\bar{\boldsymbol{g}}(\boldsymbol{\theta})$ is the mean image averaged over the parameters $\boldsymbol{\theta}$, $\boldsymbol{S}_{\boldsymbol{g}|\boldsymbol{\theta}}$ is the sample covariance matrix conditioned on the parameters, and $\det(\cdot)$ is the determinant of the matrix. Instead of evaluating the covariance matrix and its inverse for every parameter, our second approximation is to use the mean of $\boldsymbol{S}_{\boldsymbol{g}|\boldsymbol{\theta}}$ averaged over all $\boldsymbol{\theta}$. To avoid the exponential term of (1.26), the common strategy is to apply the natural logarithm on both sides and ignore the term independent of the parameters $\boldsymbol{\theta}$, which leads (1.25) to become

$$ln[pr(\boldsymbol{g}|\boldsymbol{\theta})] \cong \frac{-1}{2}(\boldsymbol{g} - \bar{\boldsymbol{g}}(\boldsymbol{\theta}))^t \bar{\boldsymbol{S}}_{\boldsymbol{g}}^{-1}(\boldsymbol{g} - \bar{\boldsymbol{g}}(\boldsymbol{\theta})) + ln[pr(\boldsymbol{\theta})] \tag{1.27}$$

Thus, after ignoring the constant term of $ln[pr(\boldsymbol{\theta})]$, the scanning linear estimator that maximizes the posterior density under these approximations is equivalent to

$$\hat{\boldsymbol{\theta}}_{SL}(\boldsymbol{g}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \Big\{ \bar{\boldsymbol{g}}^t(\boldsymbol{\theta}) \bar{\boldsymbol{S}}_{\boldsymbol{g}}^{-1} \boldsymbol{g} - \frac{1}{2} \bar{\boldsymbol{g}}^t(\boldsymbol{\theta}) \bar{\boldsymbol{S}}_{\boldsymbol{g}}^{-1} \bar{\boldsymbol{g}}(\boldsymbol{\theta}) \Big\} \tag{1.28}$$

Similarly, the distribution of noise is unknown and the training data is in generally not equal to testing data, the final equation becomes

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{SL}(\boldsymbol{g}_{tester}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \Big\{ & \bar{\boldsymbol{g}}^t_{trainer}(\boldsymbol{\theta}) \bar{\boldsymbol{S}}^{-1}_{\boldsymbol{g}_{trainer}} \boldsymbol{g}_{tester} \\ & - \frac{1}{2} \bar{\boldsymbol{g}}^t_{trainer}(\boldsymbol{\theta}) \bar{\boldsymbol{S}}^{-1}_{\boldsymbol{g}_{trainer}} \bar{\boldsymbol{g}}_{trainer}(\boldsymbol{\theta}) \Big\} \end{aligned} \tag{1.29}$$

This observer operates on the data linearly, even though, in general, the linear template is a nonlinear function of $\boldsymbol{\theta}$. In the estimation process, the observer seeks the value of $\boldsymbol{\theta}$ that will maximize this linear operation.

HO and SLO are both very useful image quality tools in medical imaging. However, in realistic cases, the sampling covariance matrix is singular since the number of images is always much less than the number of pixels of the ROI. The finite number of sample images causes the difficulty for the implementation of the prewhitening observers to work in medical imaging. In the next section, one popular approach, channelization, will be introduced.

1.11 Image Quality Evaluation in X-Ray CT system

As mentioned above, traditional approaches such MTF and NPS are both not suitable for image quality evaluation in CT systems because the assumptions are far from reality when iterative reconstruction algorithms applied. Human observers, on the other hand, are usually more expensive both in time and cost. The mathematical model observes were chosen as candidates for this research. However, a potential impediment in our selected model observers is the inverse problem of the sample covariance because the number of images is always limited. The channel mechanism (Myers and Barrett (1987)) proposed by Myers and Barrett gives this difficult problem a solution. The channel mechanism not only solves the inverse problem but also provides researchers a way to simulate or predict human-observer performance.

1.11.1 Inverse Problem of Covariance Matrix

For detection tasks and combined detection/estimation tasks, the HO (Rolland and Barrett (1992)) and SLO (Whitaker et al. (2008)) have been successfully utilized to measure image quality. To implement these two observers, first-order and second-order statistics of the data are necessary which are typically estimated using sample images from a training dataset. For a M–dimensional ROI, the inverse of the sample covariance matrix exists only when the number of images in the training dataset is larger than M+1. In CT imaging systems, however, it is difficult to have the number of training images larger than M+1 because M is very large - even for a small ROI around a lesion. The process of channelization in which the image data is processed by a small number of channels to produce channel outputs provides a solution to this problem. The number of channels is typically much less than M. For instance, the number of channel of Dense-Difference of Gauss (DDOG) (Abbey and Barrett (2001); Abbey et al. (2002)) selected in this study is 10 so the dimension of each channelized sample becomes 10 whereas the ROI size might contain M=10^4 pixels. Thus, the minimum required training size reduces to 10+1=11 instead of 10^4+1 in this specific example. The channelization makes the channelized HO (CHO) and channelized SLO (CSLO) useful for limited dataset in CT systems.

1.11.2 Channelization and Channels Selection

The channelization process is to map the image onto the channels. This process can be expressed as

$$\boldsymbol{x} = \boldsymbol{T}\boldsymbol{g} \tag{1.30}$$

where $\boldsymbol{T}$ is the matrix of channels. As mentioned in the previous section, the minimum required training size reduces from 10^4+1 to 11.

The 10 DDOG channels employed in this study have been demonstrated to match the human observer behavior in signal–detection tasks (Abbey and Barrett (2001)). In addition to reducing the dimension of the covariance matrix, DDOG channels also has been proven to have a property of mimicking human visual system (Abbey and Barrett (2001); Abbey et al. (2002)) in the pure detection task. These channels have not been verified to match human-observer performance when used in a scanning observer such as the SLO used in this work. However, we chose to use this observer model because the channels are based on the human visual system even if the scanning mechanism is not.

The general ideas used in my research have been described in this chapter. In the following chapters, I will focus on the application of model observers on CT imaging systems. Model observers will be used for the quantitative evaluations of different image reconstruction algorithms and protocols. More practical designs for the CHO and the CSLO will then be introduced in the chapter 6.

CHAPTER 2

Image Quality Assessment Using Model Observers: Detection Tasks

Task-based measures of image quality have been developed for many decades (Barrett et al. (1998)) and have been applied extensively in nuclear-medicine imaging.(Gifford et al. (2000a); Barrett et al. (2008); Gross et al. (2003)) More recently, these methods have been used in CT imaging (Yu et al. (2013)). As mentioned in first chapter, complete descriptions of three important elements are essential in the task-based assessment of image quality: (1) the task (2) the observer, and (3) the figure of merit. A number of different techniques have been developed to reduce radiation dose in X-Ray CT. Among them, iterative reconstruction (IR) algorithm has been widely expected to be an effective dose reduction for CT. To quantitatively measure the achievable dose reduction when we use IR, we will compare task-based measures of image quality of CT images reconstructed by IR and the conventional algorithm, FBP.

2.1 Tasks and Figures of Merits

In this chapter, we consider the tasks of signal–location–known signal detection as well as signal–location–unknown detection. For these tasks, the object being imaged belongs to one of two categories (or hypotheses), H_2 or the signal–present class and H_1 or the signal–absent class. For the signal detection and localization task, images that are classified as being signal present must also have this signal localized by the observer. If the localized signal is within a threshold distance from the true location, then it is considered that the observer has made a correct decision. For the signal–detection alone task, the observer's goal is to output a test statistic that is compared to a threshold to determine whether the observer classifies the image as signal present or signal absent. Receiver operating characteristic (ROC) analysis10 is a plot of the true-positive fraction (or sensitivity) versus the false-positive fraction (one-specificity) for all possible threshold values. The area under the ROC curve (AUC) is an often used figure-of-merit quantifying observer performance. For the signal–detection and localization task, we model the observer as producing a spatial test statistic map that represents the observer's confidence that a signal is present at all locations within the image. The maximum value of this map is used as the overall test statistic for the image. The location associated with the maximum is the estimated signal location when the overall test statistic is greater than the decision threshold. This framework has been used extensively in nuclear medicine imaging (Gifford et al. (2000b); Kadrmas and Christian (2002); Das et al. (2011)). A plot of the fraction of a true-positive detection with correct localizations versus the false-positive fraction is known as a localization ROC curve (LROC) (Barrett and Myers (2003)) and the area under this curve (LAUC) is used as a figure of merit (FOM) quantifying observer performance.

2.2 Channelized Hotelling Observer

The channelized Hotelling observer (CHO) (Myers and Barrett (1987); Gallas and Barrett (2003)) are valuable tools in the objective assessment of image quality. Model observers allow for rapid computation of task-based figures of merit while avoiding complicated and expensive human psychophysical studies. The CHO template was estimated by training data. The CHO template operates on both signal–present images and signal–absent images with linear channels to generate test statistics. After characterizing the observer's performance for many different thresholds, ROC or LROC curves can be generated. A flow chart of generating ROC and LROC curves are shown in Fig. 2.1. As long as the number of channels is much less than the dimension of the image data, the CHO mitigates the computational difficulties that are inherently present with the HO. The channels can be selected to either mimic the human observer or they can be selected to approximate the full HO performance by efficiently representing the signal and noise characteristics in the channel space. The CHO template can be expressed as

$$\begin{aligned} \boldsymbol{w}_{CHO} &= \Big[\frac{1}{2}(\boldsymbol{S}_{\boldsymbol{x}_{trainer,2}} + \boldsymbol{S}_{\boldsymbol{x}_{trainer,1}})\Big]^{-1}(\bar{\boldsymbol{x}}_{trainer,2} - \bar{\boldsymbol{x}}_{trainer,1}) \\ &= \Big[\frac{1}{2}(\boldsymbol{\Sigma}_2 + \boldsymbol{\Sigma}_1)\Big]^{-1}(\bar{\boldsymbol{x}}_{trainer,2} - \bar{\boldsymbol{x}}_{trainer,1}) \\ &= \boldsymbol{S}^{-1}\bar{\boldsymbol{s}} \end{aligned} \tag{2.1}$$

where $\boldsymbol{S}_{\boldsymbol{x}_{trainer,j}}$ is the covariance of the channel outputs for the H_j hypothesis, $\bar{\boldsymbol{x}}_j^{trainer}$ is the mean of the channel outputs under the H_j hypothesis, $\boldsymbol{\Sigma}_j$ is the simplified version of the channelized covariance matrix estimated by training samples under H_j hypothesis, $\boldsymbol{S}$ is the mean of sample channelized covariance matrices, and $\bar{\boldsymbol{s}}$ is the difference of the mean of channelized training data of the two hypotheses. In this dissertation, we chose to use ten dense difference of Gaussian channels (DDOG) proposed by Abbey and Barrett (Abbey and Barrett (2001)). We also compared our results using DDOG channels with the results using Gabor channels without orientation and phase terms (Wunderlich and Noo (2008)). With Abbey's DDOG model, the radial frequency profile of the jth channel is given by

$$\boldsymbol{C}_j(\boldsymbol{\rho}) = exp\Big[-\frac{1}{2}\Big(\frac{\boldsymbol{\rho}}{Q\sigma_j}\Big)^2\Big] - exp\Big[-\frac{1}{2}\Big(\frac{\boldsymbol{\rho}}{\sigma_j}\Big)^2\Big] \tag{2.2}$$

where $\boldsymbol{\rho}$ is the spatial-frequency variable and σ_j is the standard deviation of each channel. Each of the σ_j values is defined by $\sigma_j = \sigma_0\alpha^{j-1}$ from an initial σ_0. The multiplicative factor, Q > 1, defines the bandwidth of the channel. Ten channels are used in this work. Channel parameters are the same values proposed by Abbey (Abbey and Barrett (2001)), $\sigma_0 = 0.005$, $\alpha = 1.4$, and Q = 1.67. DDOG channels have been proven that it has similar image template to the image template estimated by human observer without any assumption of the noise texture of images (Abbey et al. (2002)). These channels are also known to mimic human observer performance for signal known exactly tasks. To compare the results using Gabor channels, we ignore the orientation and phase terms of Gabor channels because of the circular symmetric shape of the signals. The spatial channel widths (Wunderlich and Noo (2008)), ω_s, are 56.48, 28.24, 14.12, 7.06, and 3.53. Yu et al. (Yu et al. (2013)) has shown that Gabor channels have similar detection performance as human observers. For signal–location–unknown tasks, our scanning CHO will not mimic human observer performance because it is known that human observers do not scan the entire image but instead focus their attention on regions of interest (Krupinski (1996)). Some work by Gifford et al., (Gifford et al. (2000a)) however, has suggested that these scanning models may work reasonably well to model humans.

2.3 Training, Testing, and Variance Estimation

The training data are used to estimate the parameters that define the model observers: the means and covariance matrices. For the CHO, this training dataset is used to estimate $\bar{\boldsymbol{s}}$ and $\boldsymbol{S}$. More training images provided more accurate estimated parameters. Kalayeh and Landgrebe suggested that the required number of samples should be at least five times the number of features (Kalayeh and Landgrebe (1983)). Once these parameters are estimated, the testing dataset is used to estimate the FOM. There are several ways to estimate the variance of the FOM. Bootstrap (Efron and Tibshirani (1993)), jackknife (Dorfman et al. (1992)), and shuffle (Dorfman et al. (1992); Fukunaga and Hayes (1989)) methods have been studied previously (Wagner et al. (1997, 1998); Biden et al. (2000); Obuchowski and Rockette (1995); Obuchowski (1995); Barrett et al. (2005)). In our study, the variance of AUC and LAUC values are estimated by a completely nonparametric and unbiased approach, referred to as the Barrett, Clarkson, Kupinski method (Gallas (2006)) or the One-Shot method (Dorfman et al. (1998)). This variance-estimation method is a variant on the common Dorfman, Berbaum, Metz technique (Roe and Metz (1997a,b)) but does not rely on resampling techniques.

2.4 Phantom, ROI, and Data Acquisition

In this work, low contrast (LC) objects embedded in a MITA IQ LCD phantom (CCT 183, The Phantom Laboratory, Salem, NY) were imaged on a GE OptimaCT660 CT system. Figure 2.2 provides the head and body phantom MITA

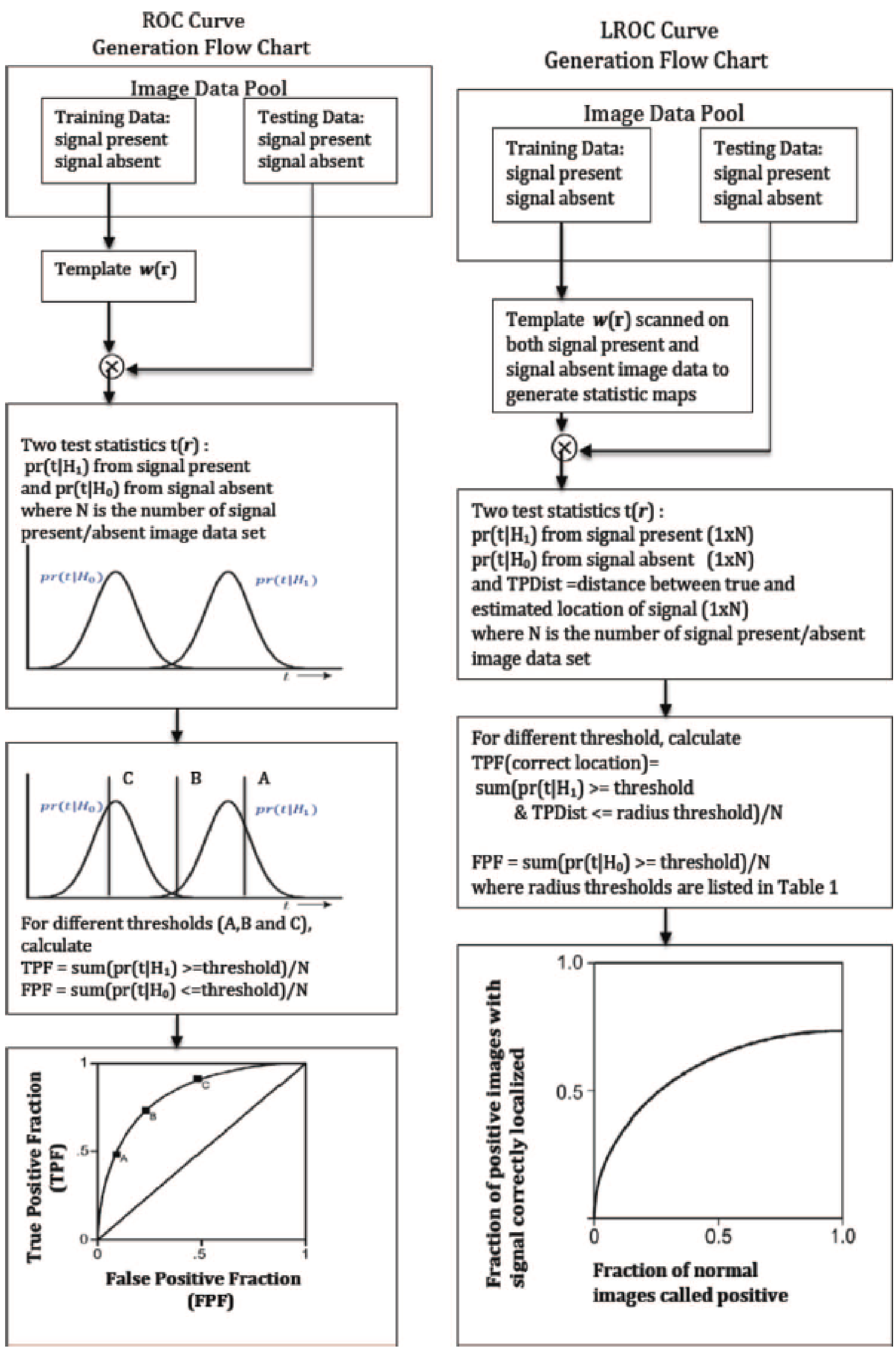

Figure 2.1: (Left) ROC study. (Right) LROC study.

phantom used in this study. The diameter of the head mode is 200 mm and the dimension of oval ring of the body phantom is 350 $\times$ 250 mm. Multiple LC objects with various contrast and size levels present in this phantom were studied. We used axial scans with rotation time = 1 s. The scan FOV (SFOV) was small body for head phantom and large body for body phantom. The slice thickness was 0.625 mm and the detector coverage was 20 mm in our study. In order to gain statistical confidence in the detectability measurement, multiple scans were required. 21 and 34 different radiation dose protocols were used for the head phantom and body phantom, respectively. For each dose protocol, 50 identical scans were acquired. A total of ten individual signal present ROI were extracted from different longitudinal locations (along the CT system table direction) from each scan. For each LC signal, signal absent ROIs were randomly extracted from two different locations to increase the variability of signal absent case (Fig. 2.3). Ten signal–absent images were extracted from one scan. The data were acquired at 120 kVp. The x-ray current was varied to achieve different exposure levels and, hence, different doses. The 50 scans and ten extracted ROI pairs per scan resulted in 500 individual ROI pairs for each LC object at every dose level. Images were reconstructed at a field of view of 180 mm with a matrix size of 512 $\times$ 512 image pixels. The order of all images was randomized before being split into training and testing datasets. The same randomized sequence was used for every study. The sizes of the ROIs varied with the size of the LC signals (see Table 2.1). The signal was always in the center of the ROI and the noise ROIs were extracted from regions far from the signals. The same set of ROIs were used both for location known and location unknown model observer studies. For the rest of this paper, we will refer to the extracted ROIs as “images”. All four different LC objects (shown in Table I) imbedded in this MITA phantom were used in this study. The center of each signal was determined by analyzing the mean image at the highest dose level (Fig. 2.3). For each object and dose level selected, out of the 500 independent signal–present and signal–absent image pairs: 200 pairs were used for training and the remaining 300 pairs for testing. ROC curves for location known exactly task and LROC curves for location unknown task were

Table 2.1: MITA phantom for detection task

Object	Diameter (mm)	Contrast (HU)	Background (HU)	ROI (mm)	Radius Threshold (mm)
1	3	14	40	35.2	3.5
2	5	7	40	35.2	3.5
3	7	5	40	35.2	3.5
4	10	3	40	45	3.5

generated and the corresponding areas under curves were calculated. As mentioned above, the variances of estimated AUC values and LAUC values were estimated via the One-Shot method (Gallas (2006)).

2.5 Detection Task Design

In this work, we consider both the task of detecting a signal at a known location and detecting a signal at an unknown location and estimating its location. For the location known detection task, CHO templates were estimated using the training dataset and applied on each test image. The resulting test statistics from the testing dataset were used to generate ROC curves. For the location unknown study, instead of placing the LC object at random locations in the image, we had LC object centered in the ROI and the model observer was applied to scan all possible locations in every test image. The CHO template for the location unknown study (shown in Fig. 2.4(b)) was generated in a similar way as the one in location known study (shown in Fig. 2.4(a)). The LROC template was created by cropping the nonzero region from the ROC template (35.16×35.16 mm or 100×100 pixel). We ignore the regions of the template with small values when we cropped to ensure that just the central feature of the template is used for scanning. Note that the larger the LROC template, the less possible locations for observer to scan. The number of possible locations affects the difficulty of task by changing the blind guess line for LROC studies as described below. In this work, the sizes of LROC template are 16.88 × 16.88 mm (48 × 48 pixel) and 26.72 × 26.72 mm (76 × 76 pixel) for the first three objects and fourth object, respectively, and the LROC template was applied at all the possible locations in each image and a test statistic was produced at these locations. A test statistic map (Fig. 2.5) was generated which represents the likelihood of a signal being present as a function of signal location. The location of the maximum test statistic was determined to be the signal location. When the distance between signal location estimated by the scanning observer and the true location (center of the statistic map) is less than a certain radius threshold, the estimated location is counted correct. The values of radius thresholds depend on the sizes of ROIs and were chosen to maintain an equal guessing curve (with area about 0.05) for all cases (Table 2.1). The blind guess line shown in all subsequent LROC curves were calculated based on the probability that a randomly selected location is counted as a correct.

2.6 Selection of Proper Ranges of AUC and LAUC Values

The AUC value was used as the figure of merit for the location known detection tasks in our study. The guessing line, which represents the worst possible ROC curve is a straight line from (0,0) to (1,1) in ROC space. Thus, the AUC value of 0.5 implies that the task is too hard. Similarly for LAUC, if the value is close to guessing line from (0,0) to (1,0.1) in our study, then this location unknown detection task is considered too difficult. On the other hand, if the AUC value or LAUC value is close to 1.0, then the task is considered too easy (trivial) for observers. In our study, various dose levels were scanned. Some of them provided AUC and LAUC values either too difficult or too easy. We restricted our analysis using the AUC values from 0.55 to 0.92 range and LAUC values from 0.1 to 0.9 range.

2.7 Measurements and Analyses

In this section, comparisons are made between the performances of FBP–reconstructed images at high and low exposure levels to that of the new IR images at low exposure levels. As mentioned above, the tasks considered were the signal–

location–known detection and signal–location–unknown detection of low contrast lesions. For the illustration of the purpose, the 5 mm in diameter and 7 HU in contrast relative to the background imbedded in the head mode at 9.31 mGy and body mode at 18.62 mGy are shown in Fig. 2.6. Because the contrast of the signal is very low, 20 FBP image slices of this object is averaged and displayed here. All the signals to be detected are considered very weak; this enables a reasonable dynamic range for comparing performances across algorithms and exposure levels.

2.7.1 Location Known Detection Task

AUC was used as the figure of merit for the location known detection tasks. The larger the AUC value is, the better the performance of the algorithm is. The guessing line, which represents the worst possible ROC curve is a straight line from (0,0) to (1,1) in ROC space. Thus, the AUC is 0.5 for a guessing observer. In this study, different dose levels were selected for comparison for every object. The dose levels were selected to ensure that the AUC values fall into the range of 0.55-0.92. In this range, the detection task is not too easy for the observer which would result in AUC values of 1, and it is also not too difficult which would result in ROC curves near the guessing line. The high- and low-AUC regions should generally be avoided when trying to make comparisons. The AUC values analyzed for the task of detecting the various signals showed that image quality of the new IR algorithm at low exposure levels is slightly better than that of FBP at high exposure levels both in head mode and body mode and the differences between the AUCs of low-dose IR and low-dose FBP are statistically significant ($p < 0.01$) for both head and body mode. Figures 2.7 and 2.8 show the AUC values analyzed by DDOG and associated error bars in head mode and body mode, respectively, for all objects. These results suggested that the dose reduction ranges from 50% to 67% and 68% to 82% for head phantom and body phantom, respectively. Same testing has also been conducted using the Gabor channels (Yu et al. (2013)), which suggests that the dose reduction fall into the range of 60%-75% and 70%-80% for head phantom and body phantom, respectively. The results from these two channels both imply that the IR images only acquired

Table 2.2: Summary of Dose Reduction

	AUC(Head)	AUC(Body)	LAUC(Head)	LAUC(Body)
DDOG	50%-67%	68%-82%	67%-75%	67%-77%
Gabor	60%-75%	70%-80%	67%-70%	67%-75%

at a small portion of dose results in the same or even better detectability than the full dose FBP data for the detection task. The summary of dose reduction in the detection task is shown in Table 2.2. Two examples of ROC curves are shown in Fig. 2.9. The blue curves, the red curves, and the green curves represent ROC curve for high dose FBP, low dose FBP, and low dose IR, respectively. Note that the ROC curves are not smooth because the number of testing data is finite.

2.7.2 Location Unknown Detection Task

Similar to the concept of AUC, the larger the LAUC value is, the better the observer performance and, hence, better image quality. The observer scanned at all the possible locations for both signal–present and signal–absent images to generate the test statistic maps. The maximum value on the map of each class was contribute to the related test statistic. The location associated to the maximum value on the map was compared to the true location of signal for the signal–present class. If the relative distance between signal location estimated by the scanning observer and the true location is less than certain radius threshold, the estimated location is counted correct. The LROC curve was then generated based on varying the statistic threshold decided by the observer and the estimated location of the signal. The dose levels in LROC studies were selected such that the LAUC values are at least twice the blind guessing value (LAUC– 0.05) and less than the value of trivial task (LAUC $= 1$). In this range, we can ensure that the task is not too easy for an observer to detect the signal and find its location correctly every time or too difficult to return the blind guess line all the time. Similar to the ROC study, dose levels may be different in each case to ensure reasonable LAUC values. LAUC values analyzed by DDOG channels and their associated error bars in head mode and body mode, respectively, for all objects are shown in Figs. 2.10 and 2.11. The differences between the LAUCs of low-dose IR and low-dose FBP are statistically significant ($p < 0.01$) for both head and body mode. With DDOG channels, the dose reduction ranges were from 67% to 75% and 67% to 77% for head phantom and body phantom, respectively. We also analyzed LAUC using Gabor channels and the dose reduction ranges were 67%-70% and 67%-75% for the head phantom and body phantom, respectively. Results from the two channels gave similar dose reduction in both head phantom and body phantom case. Again, this suggests that IR images only need a small portion of the dose to reach the same or better detectability than the traditional full dose FBP images in this study. More detail about dose reduction for each object is summarized in Table 2.2. Two examples of LROC curves are shown in Fig. 2.12. Similar to ROC curves, the LROC curves are not smooth due to the limited number of testing data. The solid blue curves, the dashed red curves, the solid green curves, and the dashed-dot black curves represent LROC curve of high dose FBP, low dose FBP, low dose new IR, and guessing line, respectively.

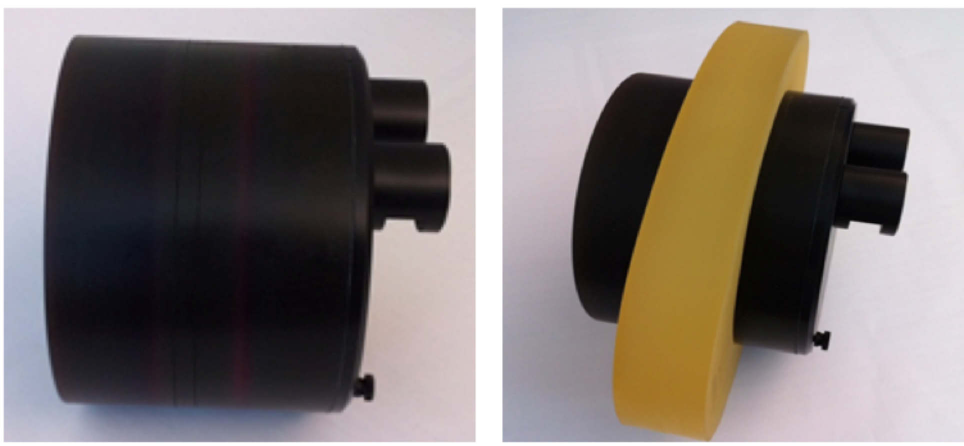

Figure 2.2: (Left) Head mode. (Right) Body mode.

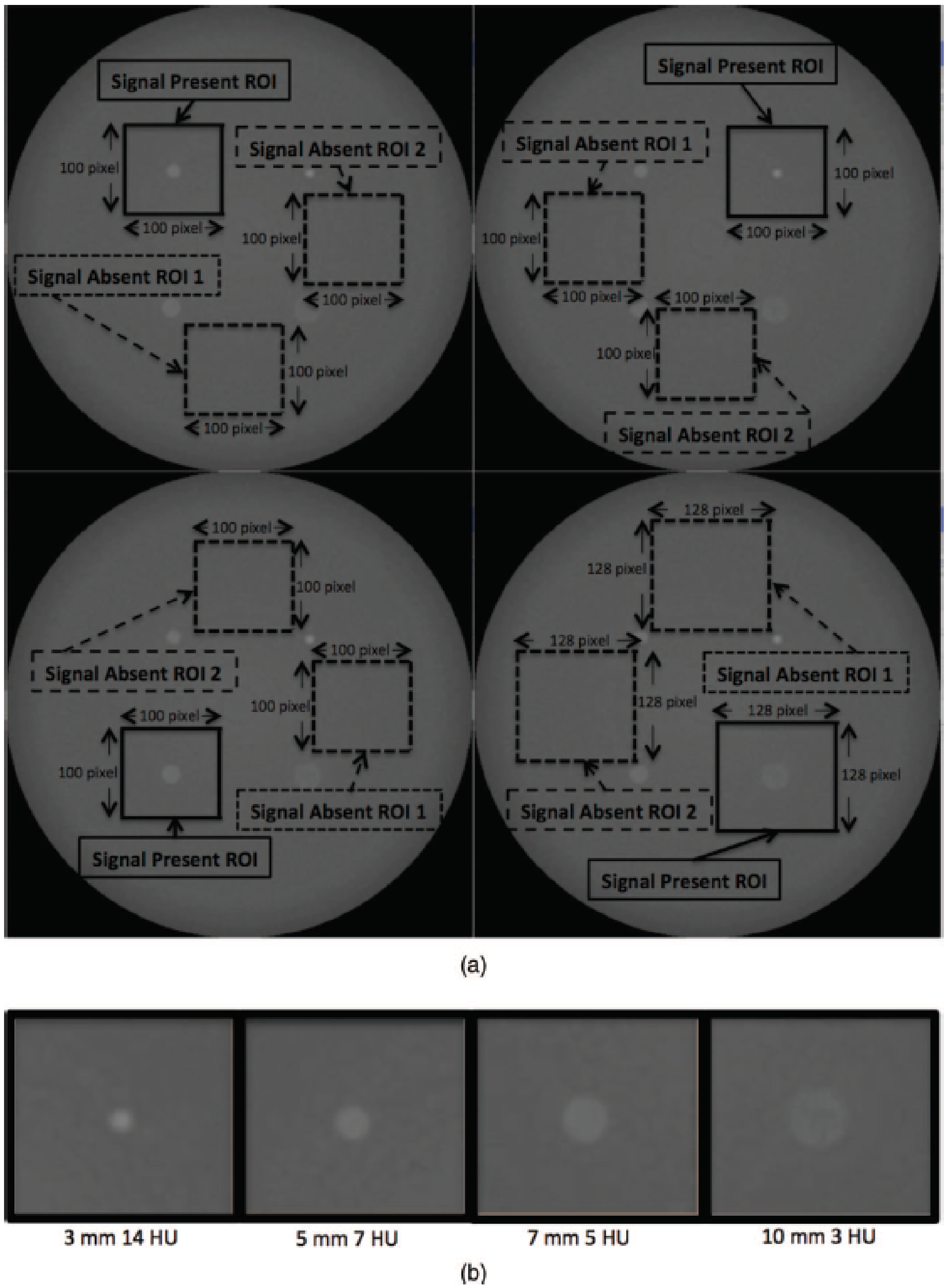

Figure 2.3: (a) Locations of signal present ROIs and signal absent ROIs for all objects. (b) Signal present ROIs for all objects embedded in the phantom. The image is the average of 200 FBP images acquired at $CTDI_{vol} = 55.86$ mGy.

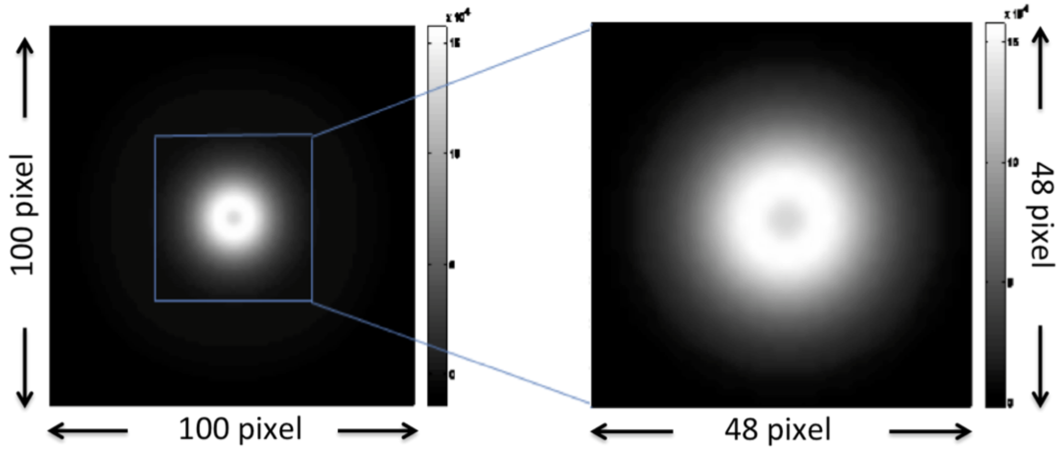

Figure 2.4: (a) Original CHO template of object 7 mm 5 HU (ROC template). (b) LROC template (Truncated ROC template). LROC template ignores the regions of small values but contains the same central feature of ROC template.

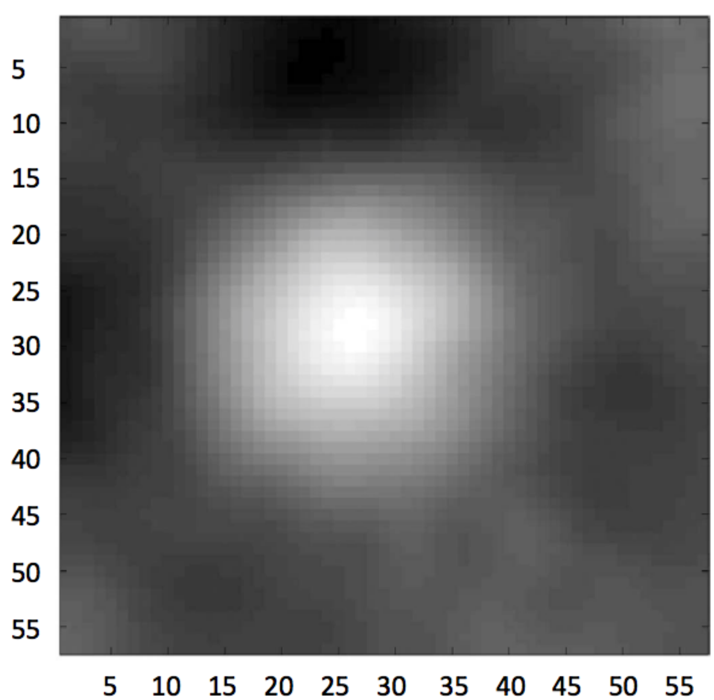

Figure 2.5: Statistic map generated by CHO LROC template scanning on one test image. The location of maximum value will be the estimated location of signal.

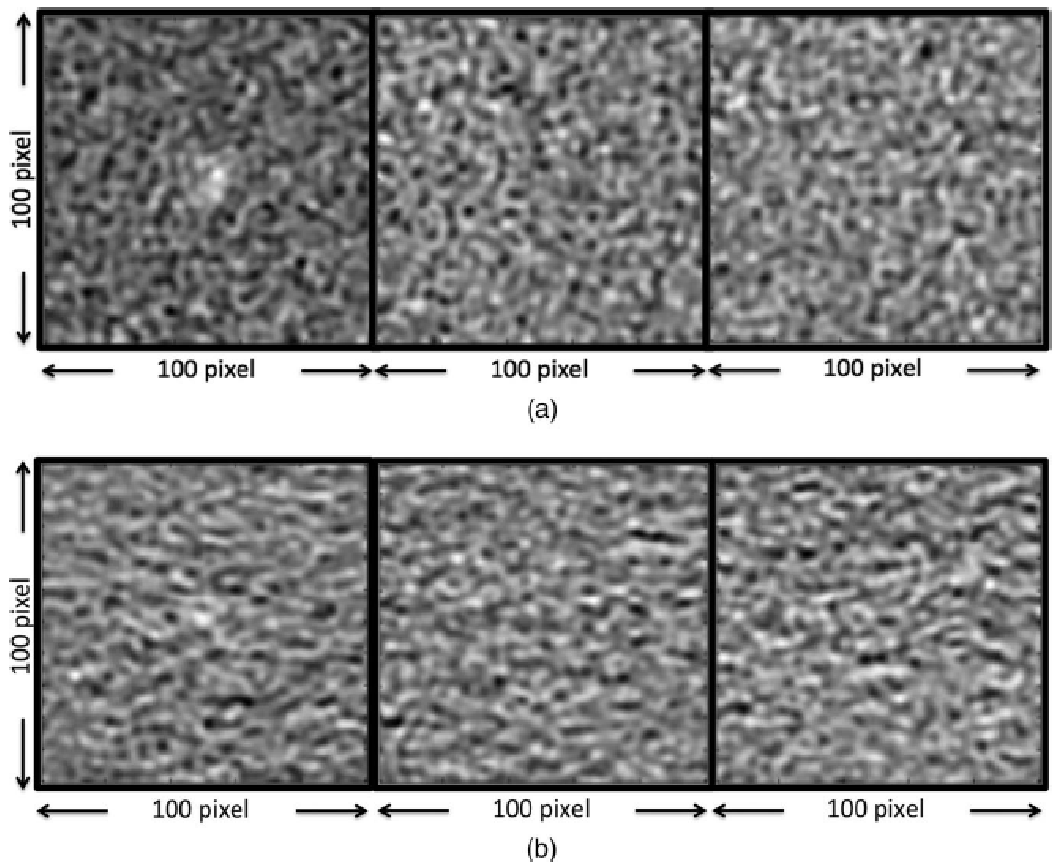

Figure 2.6: FBP CT images of object 5 mm 7 HU. (a) Head mode at 9.31 mGy. Left: signal present ROI; middle: signal absent ROI 1; right: signal absent ROI 2. (b) Body mode at 18.62 mGy. Left: signal present ROI; middle: signal absent ROI 1; right: signal absent ROI 2.

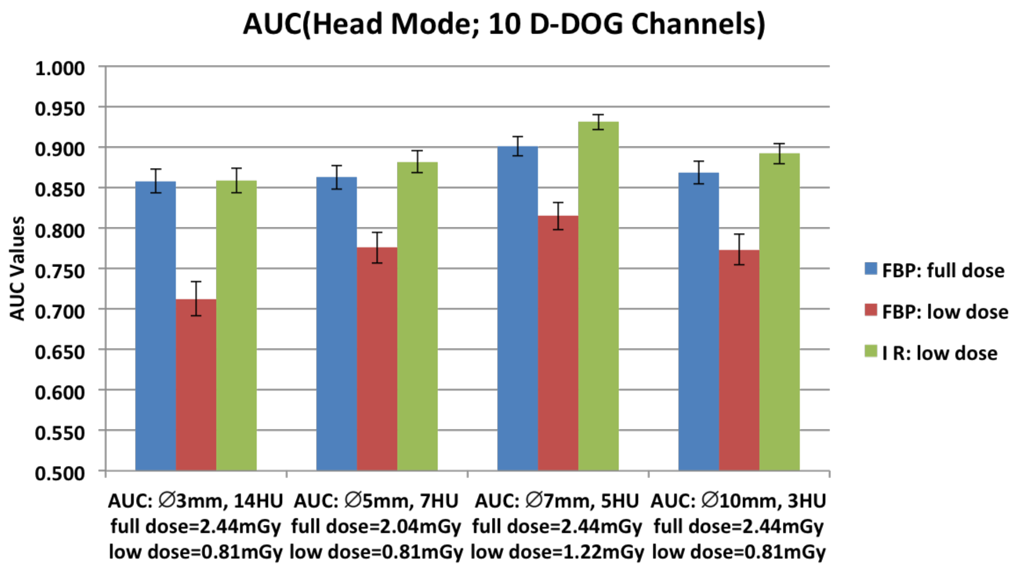

Figure 2.7: Detection task (Head mode; DDOG channels): AUC values and associated error bars for all objects.

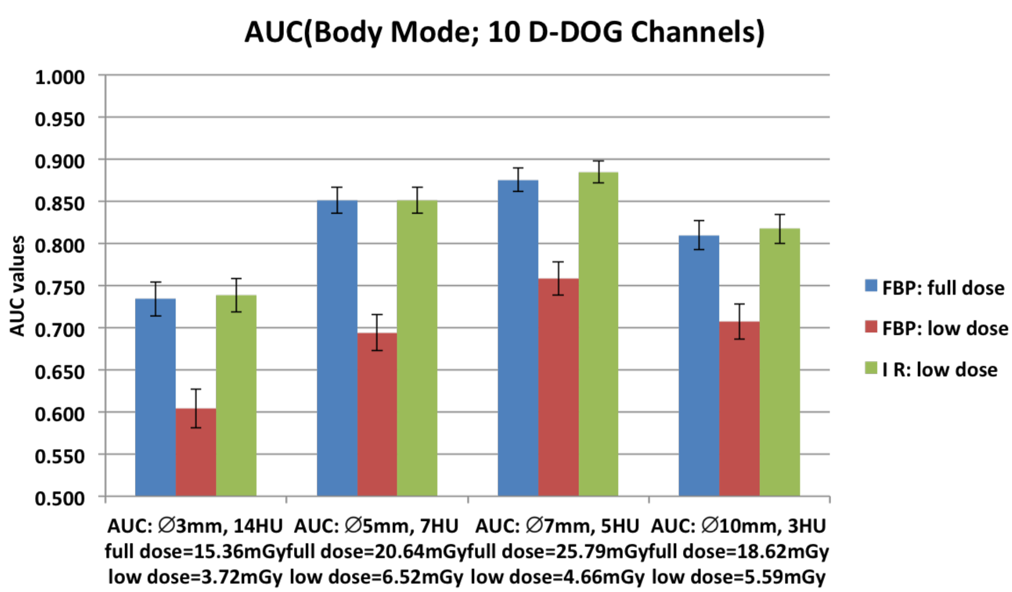

Figure 2.8: Detection task (Body mode; DDOG channels): AUC values and associated error bars.

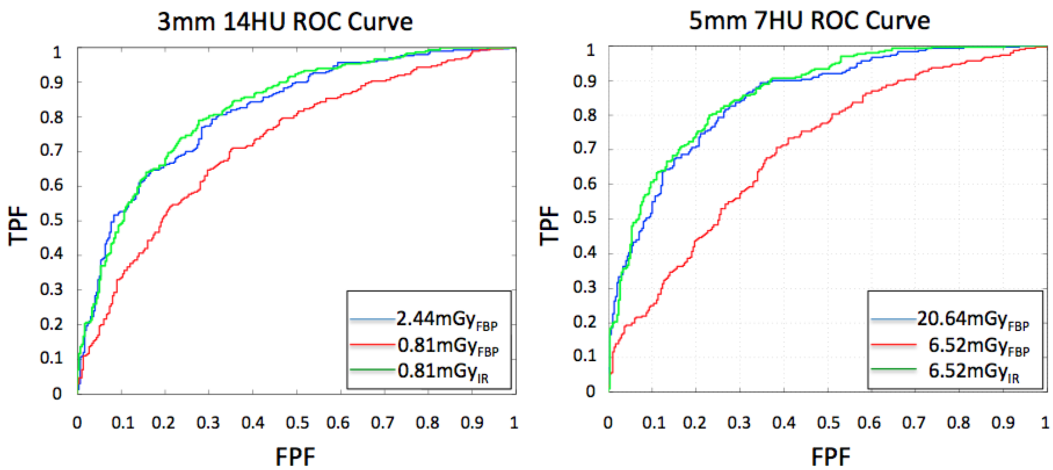

Figure 2.9: ROC curve examples. Left: 3 mm 14 HU object in head mode. Right: 5 mm 7 HU object in body mode.

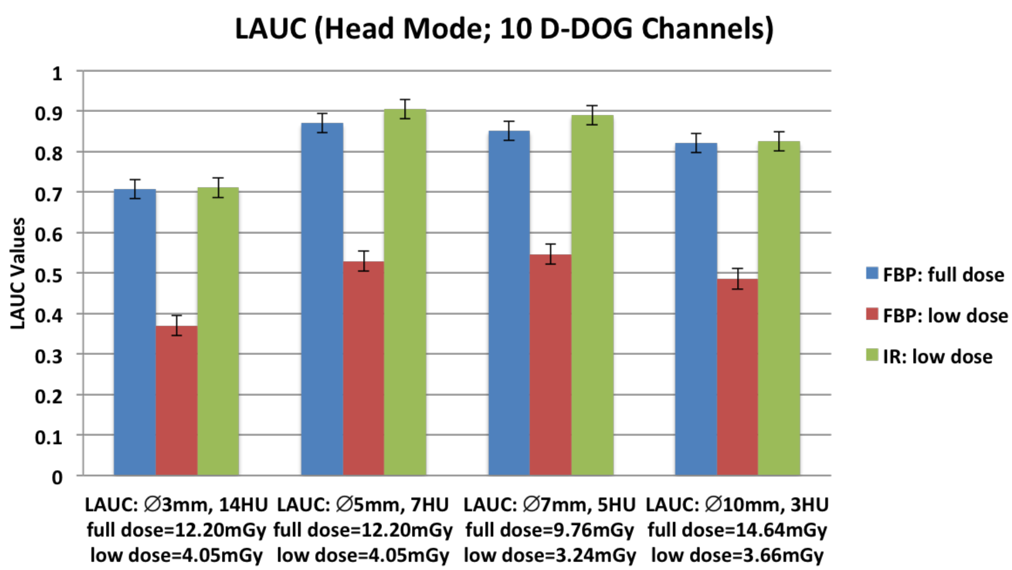

Figure 2.10: Location unknown detection task (Head mode; DDOG channels): LAUC values and associated error bars.

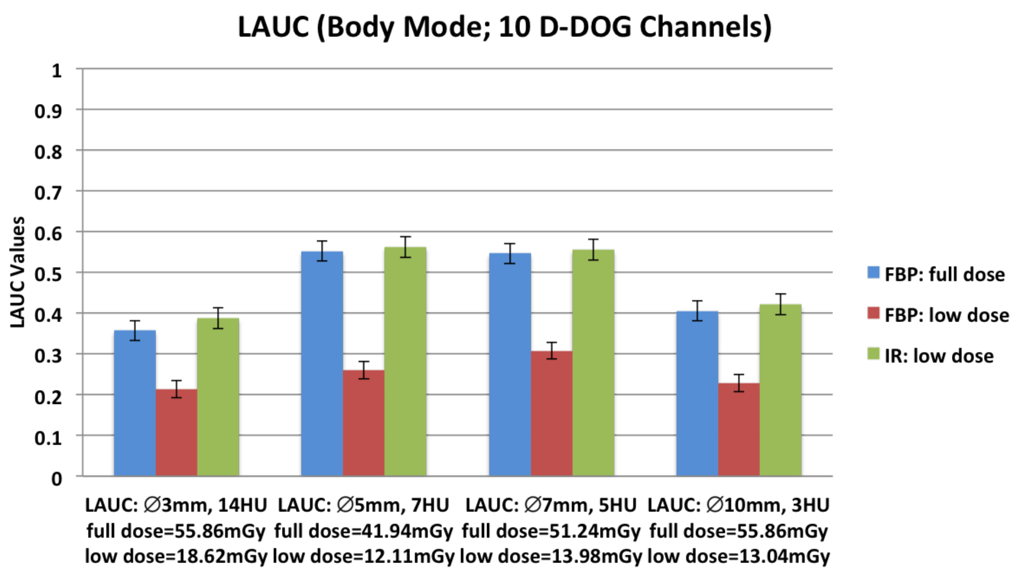

Figure 2.11: Location unknown detection task (Body mode; DDOG channels): LAUC values and associated error bars.

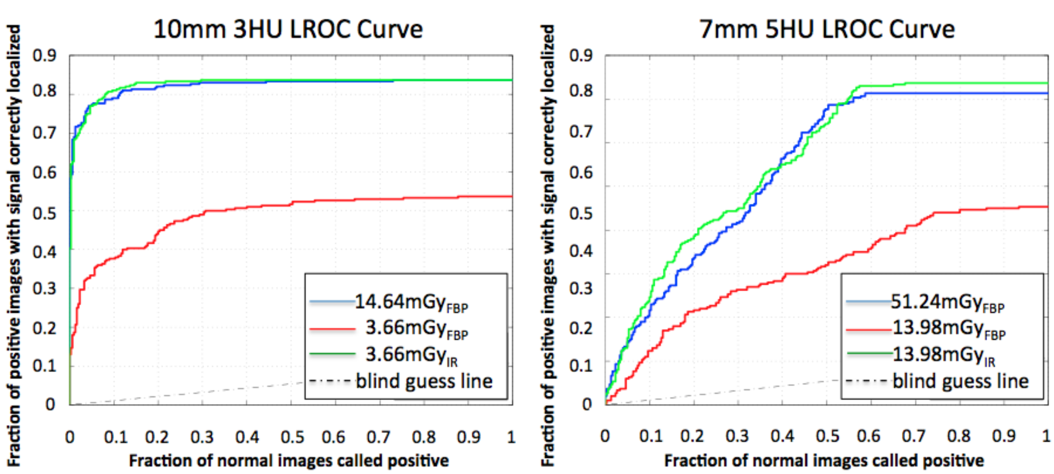

Figure 2.12: LROC curve examples. Left: 10 mm 3 HU object in head mode. Right: 7 mm 5 HU object in body mode.

CHAPTER 3

Image Quality Assessment Using Model Observers: Combination of Detection and Estimation Tasks

In chapter 2, an IR algorithm has been shown to have image quality comparable to that of traditional algorithm FBP, but with significant reduced radiation dose. 82% dose reduction is achievable as the IR algorithm is applied in CT systems. However, the dose-saving capabilities are different for different clinical tasks. In this chapter, the channelized scanning linear observer (CSLO) is applied to study the combination of detection and estimation task performance using CT image data. I will focus on the application of the CSLO on CT images to quantitatively evaluate the performance of different reconstruction algorithms under a fuller picture of tasks that involve both detection and estimation components.

3.1 Model Observer-Channelized Scanning Linear Observer (CSLO)

The concept of scanning linear (SLO) estimation (Whitaker et al. (2008)) has been illustrated in the general introduction of first chapter, and the equation is given by

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{SL}(\boldsymbol{g}_{tester}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \Big\{ & \bar{\boldsymbol{g}}^{t}_{trainer,2}(\boldsymbol{\theta}) \bar{\boldsymbol{S}}^{-1}_{\boldsymbol{g}_{trainer,2}} \boldsymbol{g}_{tester} \\ & - \frac{1}{2} \bar{\boldsymbol{g}}^{t}_{trainer,2}(\boldsymbol{\theta}) \bar{\boldsymbol{S}}^{-1}_{\boldsymbol{g}_{trainer,2}} \bar{\boldsymbol{g}}_{trainer,2}(\boldsymbol{\theta}) \Big\}\end{aligned} \tag{3.1}$$

Again, for the inverse of the sampling covariance matrix to exist with limited available image data, a channelization process is needed in our selected observer. After channelization, the above equation becomes

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{CSL}(\boldsymbol{x}_{tester}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \Big\{ & \bar{\boldsymbol{x}}^{t}_{trainer,2}(\boldsymbol{\theta}) \bar{\boldsymbol{S}}^{-1}_{\boldsymbol{x}_{trainer,2}} \boldsymbol{x}_{tester} \\ & - \frac{1}{2} \bar{\boldsymbol{x}}^{t}_{trainer,2}(\boldsymbol{\theta}) \bar{\boldsymbol{S}}^{-1}_{\boldsymbol{x}_{trainer,2}} \bar{\boldsymbol{x}}_{trainer,2}(\boldsymbol{\theta}) \Big\}\end{aligned} \tag{3.2}$$

where $\boldsymbol{x}$, again, equals channelized data and the subscript 2 refers to the signal present image class. As mentioned in the last section of chapter 1, even though these channels have not been verified to match human-observer performance when used in a scanning observer as we were using in this work, we chose to use this observer model because the channels are based on the human visual system even if the scanning mechanism is not.

3.2 Tasks and Test Statistics

The test statistic that is used to decide whether the signal will be called present or not is written as

$$\boldsymbol{t}_{CSL} = \boldsymbol{t}(\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}_{CSL}) \tag{3.3}$$

Note here that if x_{tester} in 3.2 is a signal–present image, then the test statistic generates a random variable $\boldsymbol{t}_{CSL,2}$;on the other hand, if x_{tester} is from the signal-absent image pool, then the test statistic generates a random variable $\boldsymbol{t}_{CSL,1}$. A decision that the signal is present is made when the test statistic is greater than the test–statistic threshold; otherwise the observer decides that the signal is absent. For the combined task of the signal–location–known detection and the estimation, the channelized scanning linear observer (CSLO) scans parameters in the parameter domain to search for $\boldsymbol{\theta}$ that as in eq.3.2. When the estimated parameters $\hat{\boldsymbol{\theta}}_{CSL}$, the output of eq.3.2, are equal to the true parameters $\boldsymbol{\theta}_{true}$, the estimation is counted as being correct. For the combined task of the signal– location–unknown detection and the estimation task, the CSLO first scans the location parameters in the spatial domain using the EROC template. The EROC template is actually part of 3.2 and is given by

$$\bar{\boldsymbol{S}}^{-1}_{\bar{\boldsymbol{x}}_{trainer,2}} \bar{\boldsymbol{x}}_{trainer,2}(\boldsymbol{\theta}) \tag{3.4}$$

where subscripts and superscripts have been described above and $\boldsymbol{\theta}$ is a parameter vector including locations, sizes, and contrast of signals. Due to the proper size of ROI and the dimension of current available phantoms, instead of generating a EROC template for each location, the feasible implementation is to crop the EROC template from the original image template (Fig. 3.1). The number of EROC templates or cropped EROC templates is equal to the number of LCD objects in the task.The cropped EROC template contains most of the information of original image template but with smaller size. With this approximation, the cropped EROC template was scanned over all possible locations on every testing image with a signal located at the center. There were 41×41 locations in the signal location unknown detection task. A test statistic map (Fig. 3.2) was generated after cropped EROC template scanned a single testing image. The observer then chose the location of signal regarding to the highest pixel value on the statistics map. After the location estimation, the template scanned on the size/contrast domain for size/contrast estimation or both size and contrast estimation for the given signal in the testing image. Similar to the location estimate, a test statistic map of size/contrast or size as well as contrast estimation was generated. The estimated size/contrast was based on the maximum value on this map. If the difference between estimated location and the true location (the center) was less than the location threshold (about the radius of the smallest signal) and estimated size and contrast were equal to the true values, this combined detection and estimation was considered successful. From the varying test-statistic thresholds, the two test statistic distributions of $\boldsymbol{t}_{CSL,2}$ and $\boldsymbol{t}_{CSL,1}$ and the results of estimation, an estimation receiver operating characteristic (EROC) curve (Clarkson (2007)) can be generated. The processes are illustrated in flow charts (Fig. 3.3). The area under the EROC curve, EAUC value, is the figure of merit (FOM) used in this study. The higher the EAUC value, the better the image quality of the system for the combination of detection and estimation task.

3.3 Training, Testing, and Variance Estimation

The training data are used to estimate the parameters that define the model observers, i.e., the means and the covariance matrices. As mentioned before, although the minimum required training sample size is 11, to have more accurate estimates of these statistics, 200 samples were used for training and 300 samples were used for testing. For the CSLO, the training dataset is used to estimate $\bar{\boldsymbol{x}}_{trainer,2}$, and $\bar{\boldsymbol{S}}_{\bar{\boldsymbol{x}}_{trainer,2}}$, and the equations are given by

$$\bar{\boldsymbol{x}}_{trainer,2}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_{trainer,2,i}(\theta) \tag{3.5}$$

$$\boldsymbol{S}_{\bar{\boldsymbol{x}}_{trainer,2}(\theta)} = \sum_{i=1}^{N} \left[\boldsymbol{x}_{trainer,2,i}(\theta) - \bar{\boldsymbol{x}}_{trainer,2}(\theta)\right]\left[\boldsymbol{x}_{trainer,2,i}(\theta) - \bar{\boldsymbol{x}}_{trainer,2}(\theta)\right]^t \tag{3.6}$$

$$\bar{\boldsymbol{S}}_{\bar{\boldsymbol{x}}_{trainer,2}} = \frac{1}{\text{the number of } \theta}\sum_{\theta} \bar{\boldsymbol{S}}_{\bar{\boldsymbol{x}}_{trainer,2}}(\theta) \tag{3.7}$$

Once these parameters were estimated, the testing images were substituted into 3.2 to estimate the EAUC values.

There are different ways to estimate the variance of the EAUC, and they have been discussed in Chapter 2. Here, the variance of EAUC values is estimated by a completely nonparametric and unbiased approach, referred to as the One-Shot method (Dorfman et al. (1998)).

3.4 Phantoms, Data Acquisition and Generation

In this work, low contrast (LC) objects embedded in three different phantoms were imaged on a GE Discovery CT750 HD CT system. Two of the phantoms were from Quality Assurance in Radiology and Medicine (QRM) (QRM-LC-FD1 and QRM-LC-FD4 phantoms) and the other one was from the Medical Imaging and Technology Alliance (MITA) constructed by the Phantom Laboratory (CCT 189). Examples of signal-present images of the different phantoms are shown in Fig. 3.4, Fig. 3.5, and Fig. 3.6. Based on the different designs of each phantom, QRM-LC-FD1 was used for the task of signal detection and signal-size estimation, QRM-LC-FD4 was used for the task of signal detection and signal-contrast estimation, and the MTIA phantom was used for signal detection, and signal size and contrast estimation task. The properties of LC objects in the three phantoms are listed in Tables 3.1, 3.2 and 3.3. Only axial scans were used in this work. The large body bowtie was used for

Table 3.1: QRM-LC-FD1 phantom signal parameters

Object	Diameter (mm)	Contrast Relative to Background (HU)
1	3	-2
2	5	-2
3	7	-2
4	10	-2

QRM-LC-FD1 and QRM-LC-FD4 phantom inserted in the QRM body phantom and MITA CT IQ LCD phantom with the body ring attached. The slice thickness was 0.625 mm and the collimator aperture used was 20 mm.

In this study, the X-Ray current was varied to achieve different radiation dose levels. For each dose level, 50 identical scans were acquired. A total of ten signal-present and signal-absent ROI pairs were extracted from different longitudinal locations (along the CT system table direction) from each scan. The 50 scans and ten extracted ROI pairs per scan resulted in 500 individual ROI pairs for each LC object at every dose level. Images were reconstructed at a field of view of 180 mm with a matrix size of 512×512 image pixels. A random order of 500 images was split into training and testing image datasets. The same randomized sequence was used for every study. The center of each signal was determined by analyzing the mean image. The signal was always in the center of the ROI and the signal-absent ROIs were extracted from regions distant from the signals to avoid any overlap between signal-present and signal-absent ROIs.

For each object and dose level selected, there were 500 independent signal-present and signal-absent image pairs. They were randomly split into 200 pairs for training and 300 pairs for testing in our observer studies. EROC curves for the combined detection and estimation task were generated and the corresponding areas under curves were calculated. As mentioned above, the variances of EAUC values were estimated via the One-Shot method (Dorfman et al. (1998)).

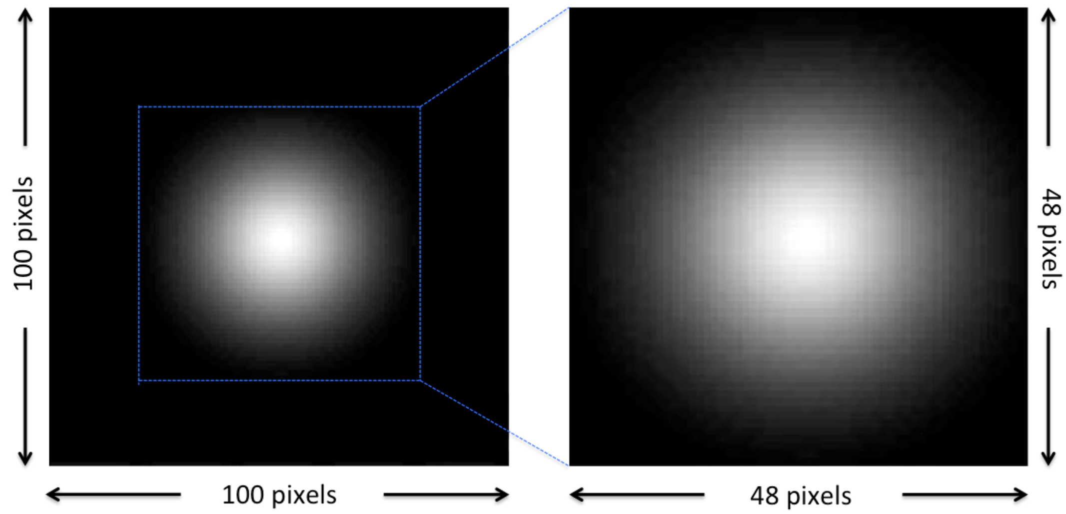

Figure 3.1: Examples of templates of one signal. Left: original EROC template. Right: cropped EROC template. The cropped EROC template is used for scanning on the spatial domain for signal location unknown task.

Table 3.2: QRM-LC-FD4 phantom signal parameters

Object	Diameter (mm)	Contrast Relative to Background (HU)
1	5	-5
2	5	-10
3	5	-25
4	5	-50

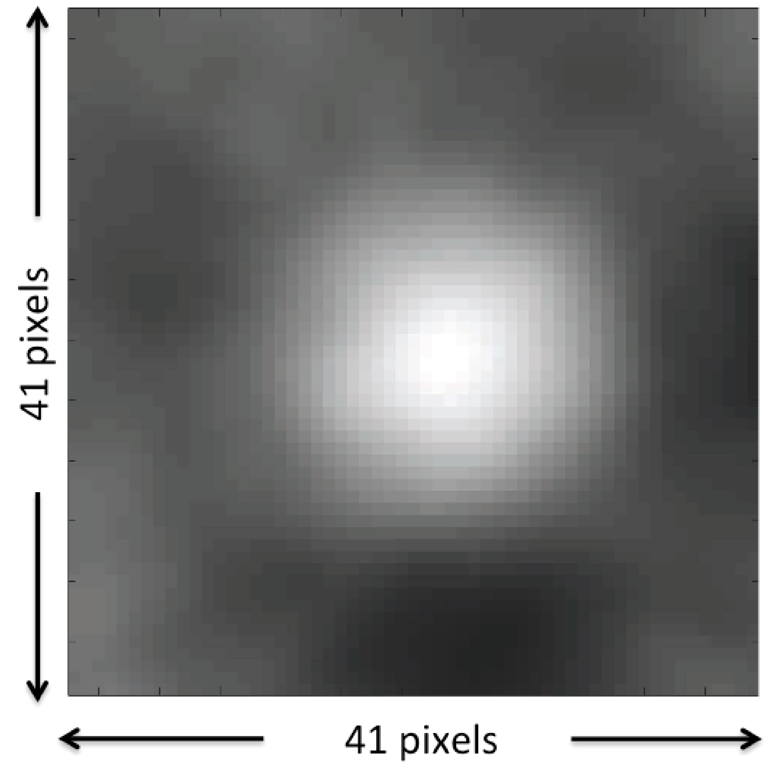

Figure 3.2: Example of a test Statistic map. After cropped EROC template of one specific signal scanned on all possible locations of a signal on a testing image, a statistics map associated to this testing signal is generated. The highest pixel value (brightest spot) in this map indicates the estimated location of the testing signal.

Table 3.3: MITA body phantom signal parameters

Object	Diameter (mm)	Contrast Relative to Background (HU)
1	3	14
2	5	7
3	7	5
4	10	3
5	15	14
6	15	7
7	15	5
8	15	3

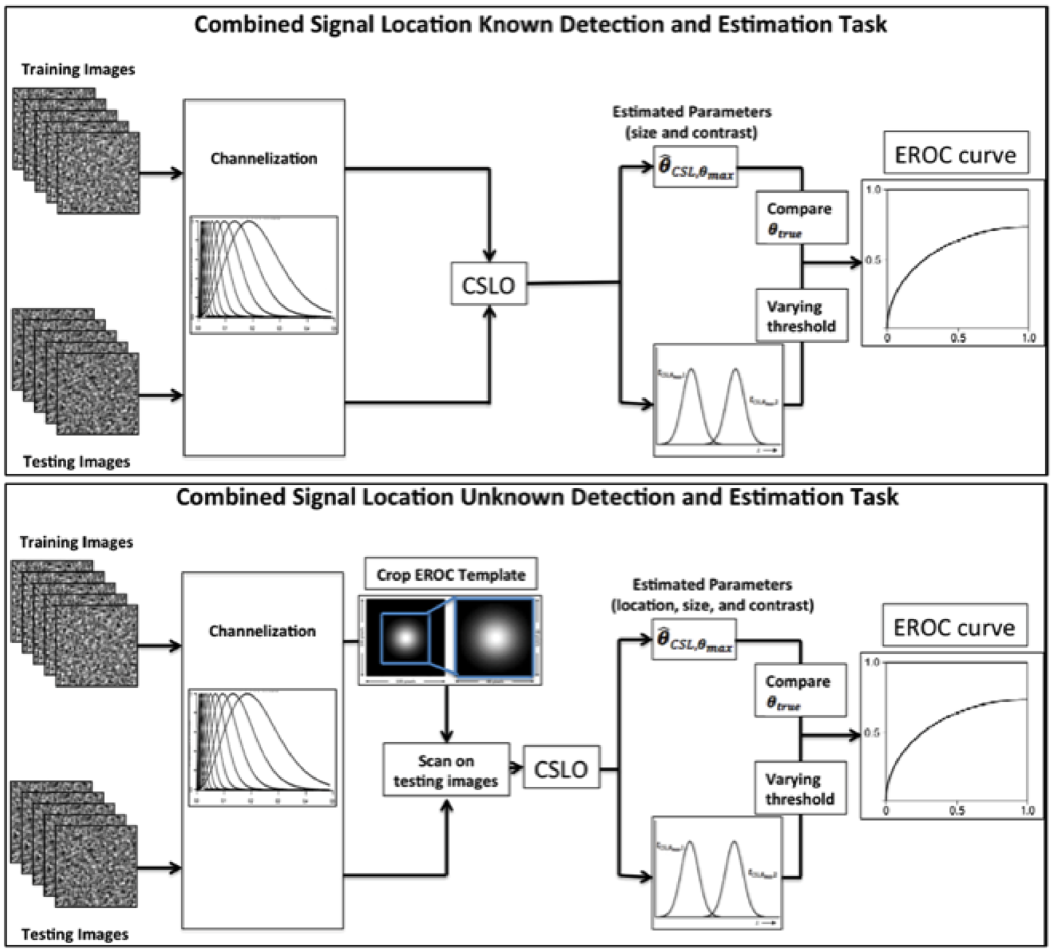

Figure 3.3: Flow charts. Top: Combined signal location known detection and estimation task. Bottom: Combined signal location unknown detection and estimation task.

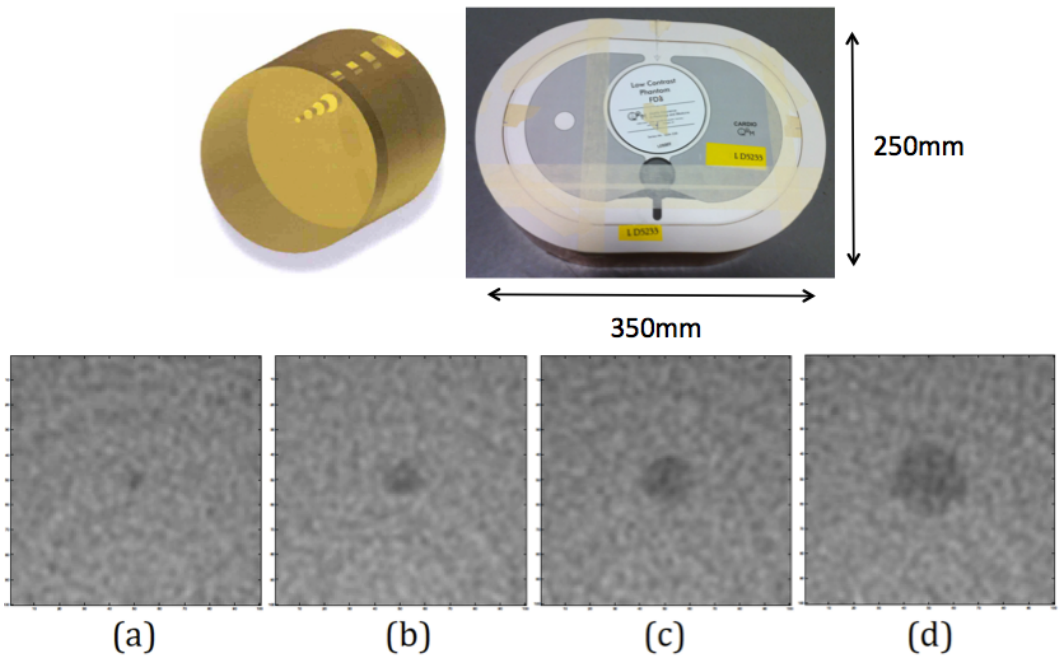

Figure 3.4: Examples of images used for the combined detection and size estimation tasks. All the images shown are the averaged result of 500 FBP images to ensure that the signals are visible in publication. The sub-images are the various signal diameters, (a) 3mm -2HU (b) 5mm -2HU (c) 7mm -2HU and (d) 10mm -2HU.

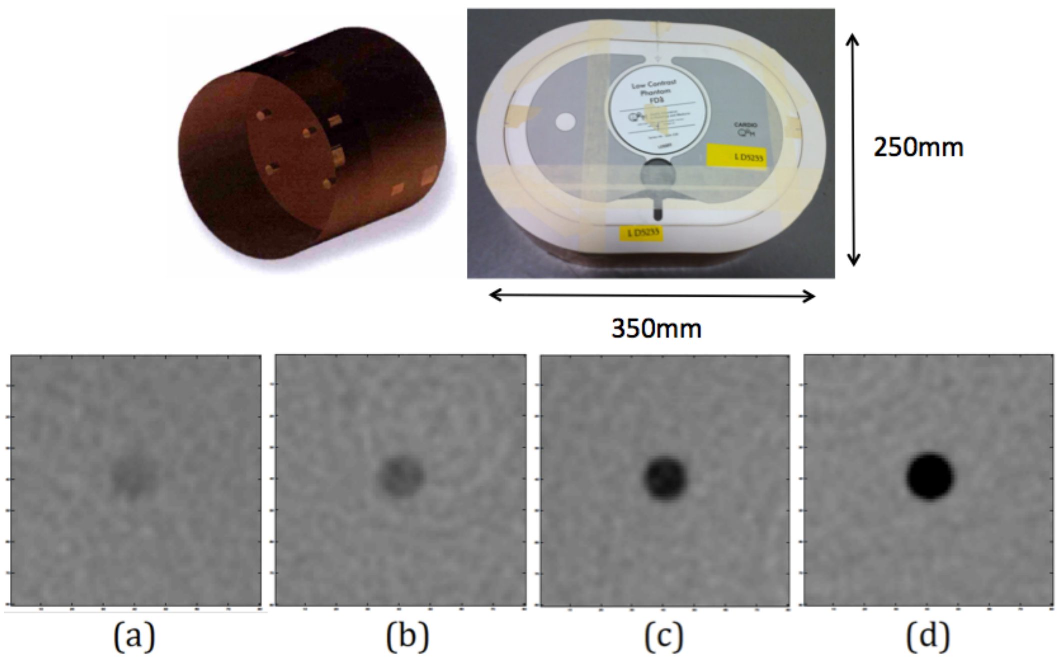

Figure 3.5: Examples of images used for combination of detection and contrast estimation task. All the images shown are the averaged result of 500 FBP images. The sub-images are the various signal contrasts, (a) 5mm -5HU (b) 5mm -10HU (c) 5mm -25HU and (d) 5mm -50HU.

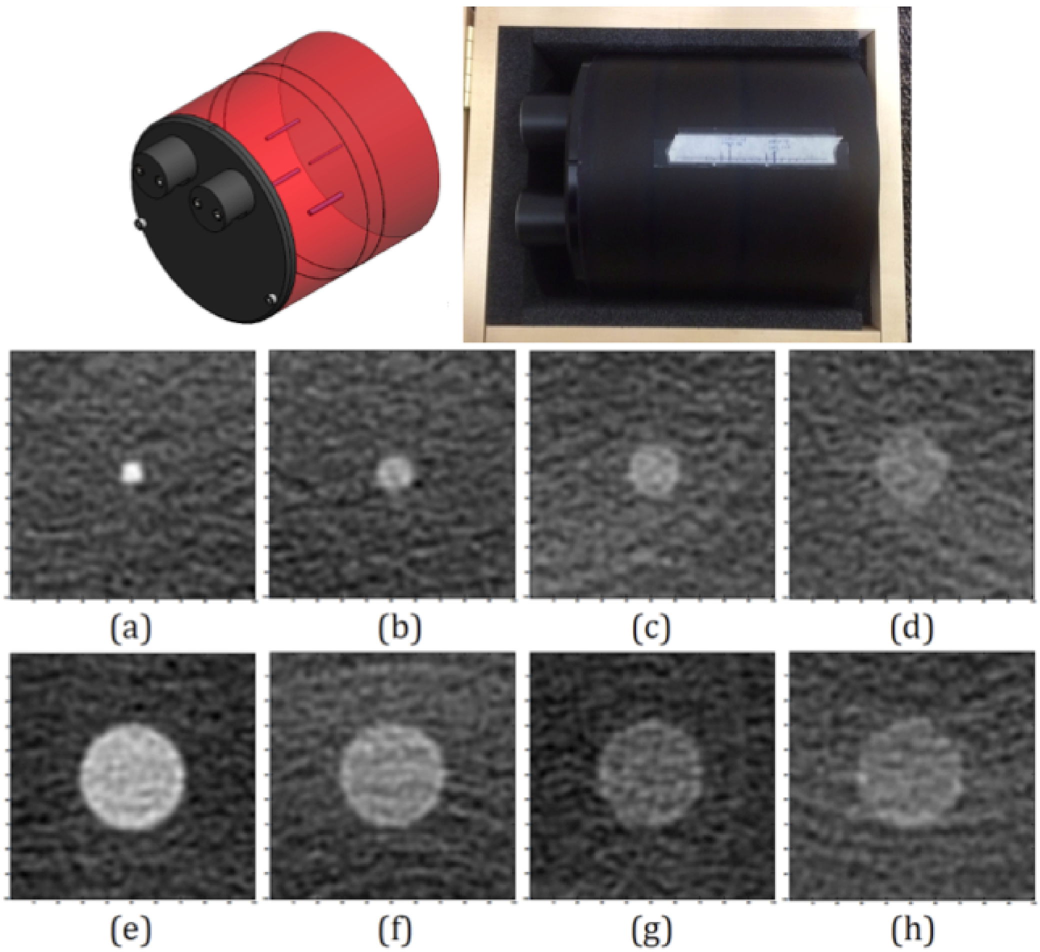

Figure 3.6: Examples of images used for combination of detection, size and contrast estimation task. All the images shown are the averaged results of 500 FBP images. The various signal parameters are: (a) 3mm 14HU (b) 5mm 7HU (c) 7mm 5HU (d) 10mm 3HU (e) 15mm 14HU (f) 15mm 7HU (g) 15mm 5HU (h) 15mm 3HU.

3.5 Measurements and Analyses

FBP and IR were compared for different tasks. Quality-Dose Characteristic (QDC) curves (Barrett et al. (2015a)) (Fig.3.7) were generated by plotting different dose levels and their associated EAUC values for both FBP and IR algorithms. For each task, 50 scans were made for each dose level and there were ten different dose levels as shown in the figure 3.7. The results show that the QDC curve of IR is higher than that of FBP at all dose levels. It suggests that the IR is better than FBP in terms of image quality at the same dose level. A comparison could also be made in the horizontal direction in QDC plot. A horizontal comparison indicates that the required dose level is lower for IR than FBP to achieve the same image quality. Based on QDC curves, achievable dose reduction can be easily derived. All phantoms were scanned axially for all experiments using the parameters mentioned above, and three major tasks (a) detection and size estimation (b) detection and contrast estimation and (c) detection, size and contrast estimation were all considered independently. For all tasks, signal-location-known detection and signal-location-unknown detection were considered.

3.5.1 The Detection and Size Estimation Task

3.5.1.1 Signal-Location-Known Detection

The QRM-LC-FD1 phantom that includes four objects with same contrast but different diameters was used in this task. The larger the area under the EROC curve (EAUC), the better the performance of the associated algorithm. The guessing line that represents the worst possible EROC curve is a straight line from (0,0) to $(1, \frac{1}{\text{the number of objects}})$. Thus, the corresponding guessing EAUC value is $\frac{1}{2} \times \frac{1}{\text{the number of objects}} = 0.125$. The factor $\frac{1}{2}$ comes from guessing the existence of the signal by an observer. The more objects, the more difficult the task gets. Due to the limited number of signals embedded in the phantoms, the exposure levels were selected to ensure that the EAUC values fall into the range of 0.125-0.9 empirically. In this range, the task was considered neither too trivial nor too difficult for the

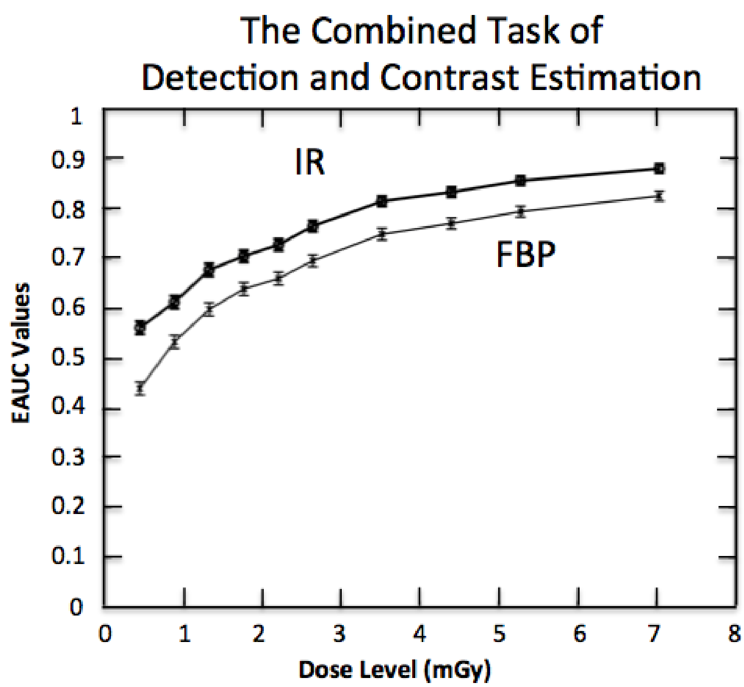

Figure 3.7: An example of Quality-Dose Characteristic (QDC) curve for the task of detection and contrast estimation.

observer to decide on the size and the existence of the signal in a testing image. Comparing two algorithms at the same image quality performance in a QDC plot, one can estimate the amount of radiation dose reduction from the IR algorithm. Results suggest that the IR algorithm required a small portion of dose but provided the same or even better detectability and estimation capability. Examples of EROC curves were shown in figure 3.8 (Left). The blue curve, the red curve, the green curve, and the dashed-dot curve represent EROC curves of the high dose FBP, the low dose FBP, the low dose IR, and the guessing line from the observer respectively. The associated EAUC values and their error bars (Fig.3.9 Left) show that image quality of the IR algorithm at low exposure levels is almost the same or even slightly better than that of FBP at high exposure levels. The dose reduction is approximately 50% based on the QDC curve. The differences between the EAUC values of low-dose IR and low-dose FBP are statistically significant ($p < 0.01$). Note that the EROC curves are not smooth because the number of testing data in a real situation is always finite.

3.5.1.2 Signal–Location–Unknown Detection

In the case of signal–location–unknown detection task, the observer had to decide whether the signal is in the provided testing image and estimate the location of the signal before estimating the size of the signal. As mentioned in the Section Methods and Materials, the location threshold is about the radius of the smallest signal. The total number of possible locations for signals is 41×41. Thus, similar to signal location known study, the guessing value of EAUC is given by

$$\frac{1}{2}\frac{\text{the area of a circle of radius equal to the location threshold}}{\text{the number of objects} \times \text{the number of possible locations}} \tag{3.8}$$

In this case the guessing value is about 0.0043. In this study, the dose reduction is 50%. The examples of EROC curves are show in Fig.3.8 (Right) and the EAUC values related to these curves are shown in the right figure of Fig.3.9. The difference between the highest EAUC and lowest EAUC in Fig.3.9 (Right) is significant ($p < 0.01$).

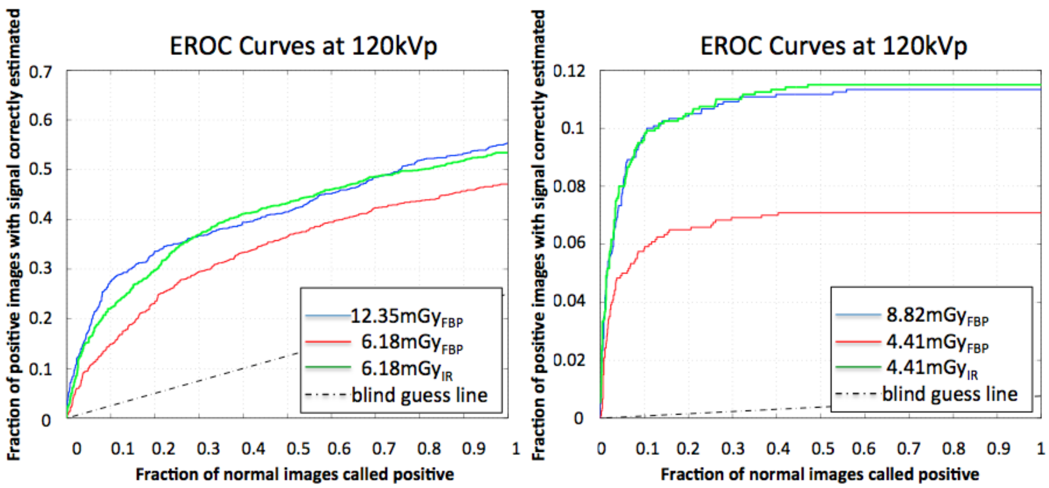

Figure 3.8: EROC curves (QRM-LC-FD1 body phantom). Left: The combination of signal location known detection and size estimation task. Right: The combination of signal location unknown detection and size estimation task.

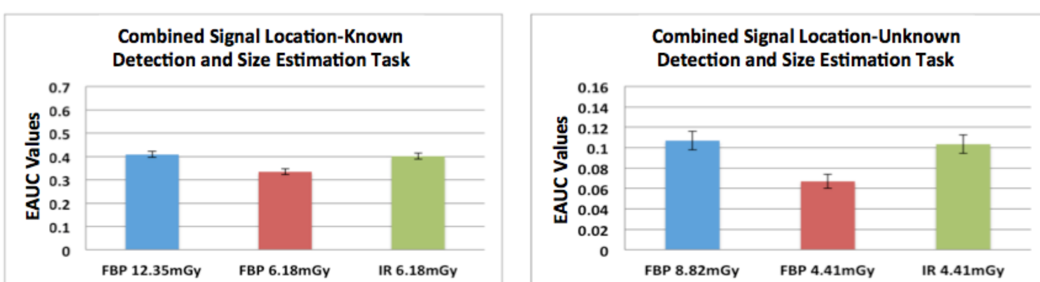

Figure 3.9: EAUC values and associated error bars of the combined task of detection(Left: signal location known; Right: signal location unknown) and size estimation regarding to Fig.8. (QRM-LC-FD1 body phantom).

3.5.2 The Detection and Contrast Estimation Task

3.5.2.1 Signal–Location–Known Detection

There are four signals with the same diameter but different contrasts embedded in the QRM-LC-FD4 phantom. The easier the observer can decide the existence of the signal from the images and recognize the correct contrast among four different contrasts, the higher the EAUC values. Since there are four objects in the phantom, the guess line is again from (0,0) to (0, $\frac{1}{4}$) and the associated EAUC value is also 0.125. The relevant EAUC range was chosen between 0.125 and 0.9 empirically. Results suggest that the dose reduction is approximately 50%. The two EROC curves achieved are shown in figure 3.10 (Left). The blue curve, the red curve, the green curve, and the dashed-dot curve represent EROC curves of the high dose FBP, the low dose FBP, the low dose IR, and the guessing line respectively. The EAUC values and their error bars are shown in figure 3.11 (Left). In this task, EAUC values of the low-dose IR and the low-dose FBP are significantly different in statistics ($p < 0.01$).

3.5.2.2 Signal–Location–Unknown Detection

In the combined task of signal–location–unknown detection and contrast estimation, the observer first scanned on every possible location of the signal on both the signal-present and signal-absent images. If the deviation between estimated location and the true location was less than the location threshold, the observer then continued to estimate the contrast of testing signal. If the estimated contrast was equal to the true contrast, the task was considered successful. In other words, the task was considered a failure if either estimated location was away from the true location or the estimated contrast was wrong. The guessing EAUC value is the about 0.0118 according to the above. The result in this study shows the dose reduction is about 50% if IR algorithm is applied. The related EROC curves and EAUC values are shown in Fig.3.10 (Right) and Fig.3.11 (Right). Again, the p value is less than 0.01 for FBP low dose case and IR low dose case.

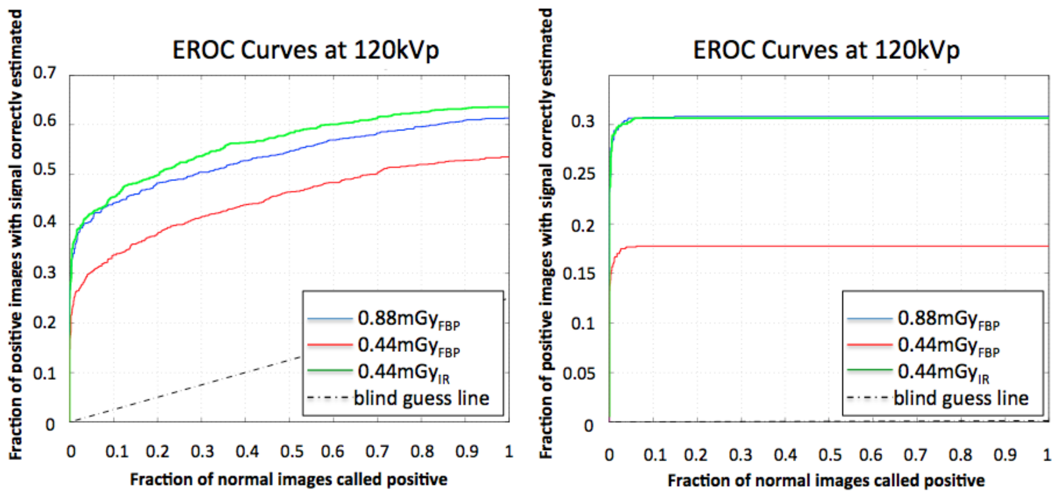

Figure 3.10: EROC curves (QRM-LC-FD4 body phantom). Left: The combination of signal location known detection and contrast estimation task. Right: The combination of signal location unknown detection and contrast estimation task.

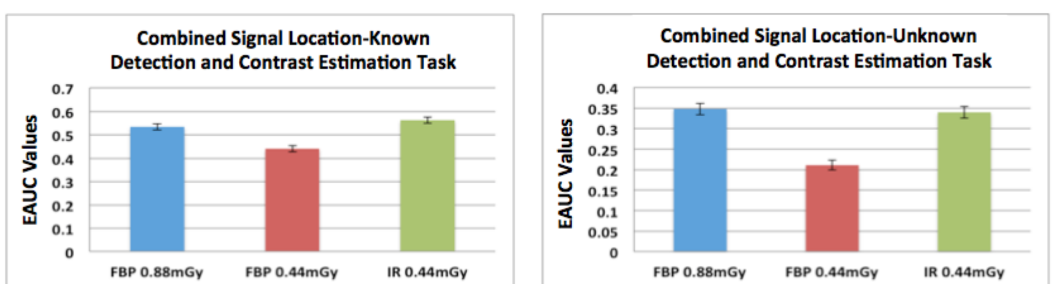

Figure 3.11: EAUC values and associated error bars of the combined task of detection (Left: signal location known; Right: signal location unknown) and contrast estimation regarding to Fig.10. (QRM-LC-FD4 body phantom).

3.5.3 The Combination of Detection, Size and Contrast Estimation Task

3.5.3.1 Signal–Location–Known Detection

In the combination task of detection, size and contrast estimation, the observer first needed to decide whether the signal was in the image and then determine the size and the contrast of the signal. In total, there are eight objects embedded in the MITA phantom. The objects 5-8 are actually designed to assist a user to measure the contrast values of the objects 1-4. In this work, we used them differently. As can be seen from Table 3.3, objects 5, 6, 7 and 8 have the same size (15mm) in diameter but different contrasts. Four pairs (object 1 and object 5; object 2 and object 6; object 3 and object 7; object 4 and object 8) have the same contrast but different diameters. The lower limit or the guessing line value is 0.0625 according to the equation mentioned in previous sections. Results indicate that the dose reduction is approximately 54%. The EROC curves are shown in figure 3.12 (Left). The EAUC values and the associated error bars are shown in figure 3.13 (Left). Comparing the low dose IR and FBP reconstructions, p value is less than 0.01. This suggests the difference between the two algorithms is significant.

3.5.3.2 Signal–Location–Unknown Detection

In this section, we consider the most complete picture of the combination task in this whole study-detecting signal, determining the location of the signal, estimating the size of the signal, and estimating the contrast of the signal. The guessing value is about 0.0021. The process of task is similar to previous tasks mentioned before so the description of process will be skipped here. The result indicates that 50% dose reduction could be reached when the IR algorithm used in the system. EROC curves are shown in the Fig.3.12 (Right). The significant difference ($p < 0.01$) between the EAUC value of IR low dose and FBP low dose is shown in Fig.3.13 (Right).

The improvement of image quality of IR is substantial. This improvement not only helps observers determine if tumors exist and where the tumors located relatively correctly but also helps observers distinguish the physical properties such as the size and the contrast of the targeted lesion more easily.

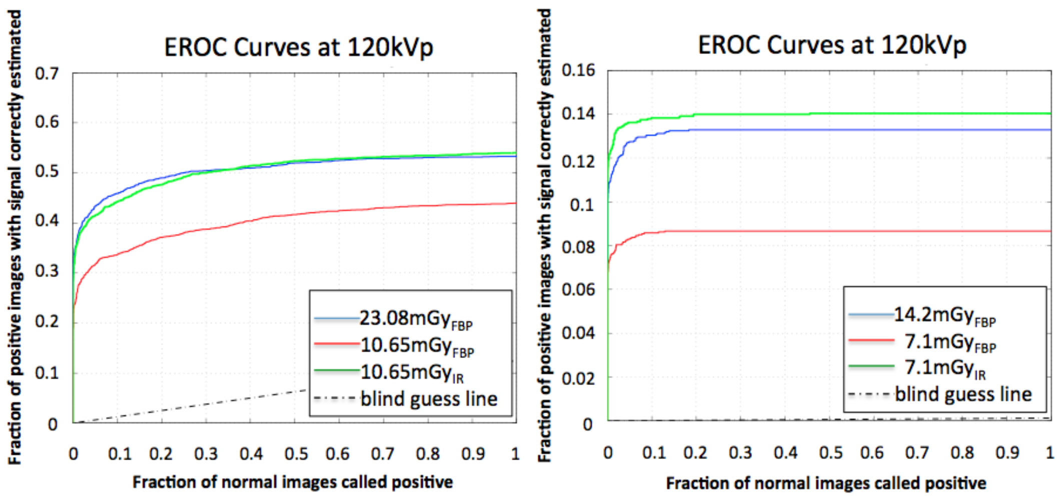

Figure 3.12: EROC curves (MITA body phantom). Left: The combination of signal location known detection, size and contrast estimation task. Right: The combination of signal location unknown detection, size and contrast estimation task.

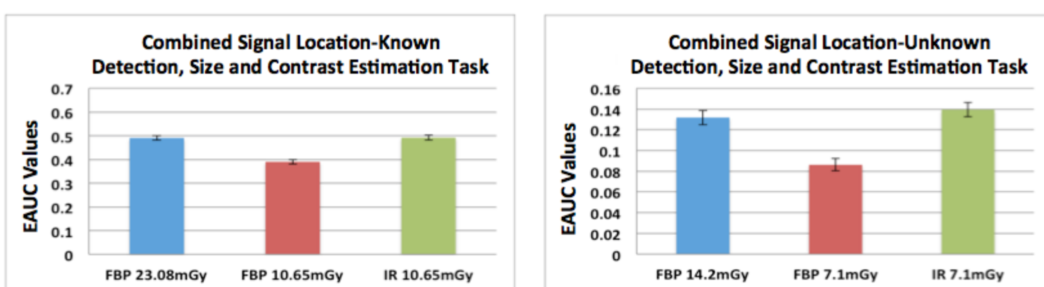

Figure 3.13: EAUC values and associated error bars of the combined task of detection (Left: signal location known; Right: signal location unknown), size and contrast estimation task regarding to Fig.12. (MITA body phantom).

CHAPTER 4

CT Protocols Optimization Using Model Observers

Advancements in multidetector-row CT technology have allowed for faster image acquisition and improved isotropic imaging. These benefits have led to a tenfold increase in the number of CT examination over the past decade. In the United States and Europe, the increase of CT studies followed the introduction of multidetector-row technology has led to significant increase in the radiation dose related to CT scanning. Although CT scanning makes up approximately 15% of radiologic examinations, it represents the largest source (70%) of medical radiation exposure. This raised the responsibility in the CT community to develop different protocols to reach the goal of as low as reasonably achievable (ALARA). Several parameters associated to radiation dose such as pitch, milliamperes (mAs), kilovolt peak (kVp), and collimation were all investigated in different protocols. In this chapter, we focused on two parameters, milliamperes and kilovolt peak. The purpose of this study is to develop an approach to optimize CT protocols based on the tasked-based image quality assessment using model observers.

4.1 Tasks and Figures of Merits

Traditionally, contrast to noise ratio (CNR) is used as an image quality metric as the energy dependent or contrast enhanced objects are scanned under different kilovolt peak. However, CNR is not in generally considered an optimal figure of merit (FOM) because of the lack of size information. Thus, in our studies, tasked-based image quality metrics using channelized Hotelling observer (CHO) (Tseng et al. (2014)), and channelized scanning linear observer (CSLO) were employed. The first one has been used for the pure detection task (Tseng et al. (2014)) in CT systems for dose reduction (Tseng et al. (2014)) evaluation. The second one was recently proposed for the dose reduction performance as well (Tseng et al. (2015)). For different tasks, different FOMs were used. For the pure detection task, the detectability generated by CHO was used as the FOM. For a pure estimation task, the correct estimation ratio (CER) produced by CSLO was used as the FOM. Finally, for the combination task of detection and estimation, the estimation receiver operator characteristic curve (EROC) was generated based on the combined outputs of the detection and estimation task by using CSLO. The FOM for the combination task was the EAUC value defined by the area under the EROC curve (Clarkson (2007)). The concepts of CHO and CSLO have been addressed in chapter 2 and 3, and the details will not be repeated here. The statistics, mean and variance, of figure of merits were estimated by 10 independent iterations. The higher the value is, the better the image quality.

4.2 Phantoms and Real Data Acquisition

In this work, iodine objects of different concentrations were inserted into a head (Fig. 4.1 (a)) mode, and a body mode (Fig. 4.1 (b)) Gammex phantom (Tissue characterization phantom; Gammex 467). Four different concentrations of iodine cylinders were inserted. Only stop-and-shoot scans were used in this work, and head and large body bowtie were used for the head mode, and the body mode study respectively. The thickness of slice was 0.625 mm and the collimator aperture used in the scan was 20 mm.

In this study, the X-Ray tube current and voltage were both varied to achieve the same $CTDI_{vol}$ value for each protocol. 4.5 mGy and 9 mGy were scanned for head and body phantoms, respectively. The window width and window level were selected at default, 400 HU and 40 HU, respectIvely. For each protocol, 20 identical scans were acquired. A total of 32 images were obtained from different longitudinal locations from each scan. The 20 scans and 32 slices from each scan resulted in 640 images. Images were reconstructed by the FBP algorithm at a field of view of 180 mm with a matrix size of 512 $\times$ 512 pixels. The signal-absent ROIs (100 pixels $\times$ 100 pixels) from 640 images were used for a better estimate of the noise power spectrum (NPS). The NPS was used for validation of our simulation data. The signal–present ROIs (100 pixels $\times$ 100 pixels) were used for design of simulated signals. In the following section, methods of generating simulation images including signal-absent and signal-present class will be illustrated.

4.3 Generation of Simulation Images

To perform an observer study correctly, the ratio of size of an object to the size of ROI is usually $\frac{1}{4}$ to $\frac{1}{5}$. Current Gammex phantoms (Fig. 4.1) were not designed ideally for this study. The insert iodine cylinder of diameter 2.8 cm was too large compared to a ROI of 3.5 cm in size. Besides, the air gaps between the inserted iodine object and phantom also caused misleading results in an observer study. To avoid these limitations, simulated images were generated based on real images for this study instead.

A simulated signal present image can be described as

$$\text{signal-absent image } \boldsymbol{g}_1 = \boldsymbol{n} + \boldsymbol{b} = \boldsymbol{HW} + \boldsymbol{b} \tag{4.1}$$

$$\text{signal-present image } \boldsymbol{g}_2 = \boldsymbol{n} + \boldsymbol{b} + \boldsymbol{s} = \boldsymbol{HW} + \boldsymbol{b} + \boldsymbol{s} \tag{4.2}$$

where the subscript index 2 means signal-present case, the $\boldsymbol{s}$ refers signal and it is circular with blurred edges. The real signal was obtained from the signal-present images that were generated by scanning the phantom with iodine objects inserted. The edge profile of simulated signal was compared to the edge profile of real signal for validation (Fig. 4.2). The term $\boldsymbol{b}$ is the phantom background and was estimated using signal-absent images that were scanned by the phantom without iodine objects inserted. The noise structure, $\boldsymbol{n}$, was simulated using correlated Gaussian noise. The convolution of a blurring filter with white Gaussian noise can be represented with matrix operator $\boldsymbol{H}$ and a zero mean white Gaussian vector $\boldsymbol{W}$. The matrix $\boldsymbol{H}$ can be estimated using real data by estimating the autocorrelation function $\boldsymbol{R_g}$

$$\boldsymbol{R_g} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{g}_{1,i}^{\text{real}} \boldsymbol{g}_{1,i}^{\text{real},t} = \langle \boldsymbol{g}_1 \boldsymbol{g}_1^t \rangle = \langle (\boldsymbol{HW} + \boldsymbol{b})(\boldsymbol{HW} + \boldsymbol{b})^t \rangle = \sigma^2 \boldsymbol{H}\boldsymbol{H}^t + \boldsymbol{b}\boldsymbol{b}^t \tag{4.3}$$

where superscript t means transport and σ is the standard deviation estimated from real image data. $\boldsymbol{R_g}$, σ, and $\boldsymbol{H}$ are all dose level dependent.

To verify that the simulated noise background is valid, the radial average noise power spectrum (NPS) profile of the simulation and the NPS profile of real image were compared (Fig. 4.3). The comparisons (Fig. 4.3) indicate the simulated noise background is very similar to the real noise background. An example of showing the similarity of a simulation image and a real image is shown in figure 4.4. A flow chart of the generation of simulation images is shown in figure 4.5.

4.4 Creation of Tasks

To perform our estimation task, 500 simulation images were generated for every object per protocol. The size of image is 100 × 100 pixels (3.5 cm × 3.5 cm). There were 250 images used to be training dataset and the other 250 images were used in the testing images. In this study, there were 4 different sizes of objects (3 mm, 5 mm, 7mm, and 10 mm in diameter). For each category, there were 5 different concentrations of iodine objects (2 mg/mL, 2.5 mg/mL, 3 mg/mL, 3.5 mg/mL, and 4 mg/mL). This led to 20 different iodine objects for our tasks. All objects were in the center of images. The dose levels were simulated at 9.5 mGy (Roberts et al. (2002)) and 13.2 mGy (Nakayama et al. (2005)) for head and body mode studies, respectively. The window width was adjusted to be 350 HU and the window level was shifted to 50 HU optimally for the head mode study (Roberts et al. (2002)). For the body mode study, the optimal window width and window level were selected at 300 HU and 50 HU, respectively (Nakayama et al. (2005)). For the section thickness, 20 mm (Roberts et al. (2002)) and 5 mm (Nakayama et al. (2005)) were used for the head and body mode study, respectively. For pure detection, detectabilities for protocols were compared for every different size and concentration of iodine objects (Fig. 4.8 and 4.9). For the pure estimation task, CER based on the contrast estimation was compared for each protocol (Fig. 4.10). For the combined task of detection and estimation, the observer determined the existence of the object and then estimated the size and contrast of the object. EAUC values were used for evaluation (Fig. 4.11). There were 10 experiments performed for each task to estimate the mean and variance.

4.5 Measurements and Analyses

4.5.1 Contrast to Noise Ratio (CNR)

The results analyzed by the traditional image quality metric, CNR, show that the 80 kVp protocol provides the best image quality with varying iodine concentration for both head mode and body case (Fig. 4.6 and Fig. 4.7). The results also indicate that the difference of CNR between each protocol increases as the concentration increases. However, the CNR value does not change with varying size because this metric does not extract any information about sizes.

4.5.2 Detection Task: Detectability

Unlike the results from CNR, the results from the pure detection task of the head mode case (Fig. 4.8) indicate that the difference between 80 kVp and 100 kVp is much smaller when the size of object is small. The results from the body mode under the pure detection task (Fig. 4.9) are very different from the results from CNR. The detectability of 80 kVp and 100 kVp is about the same for most of the cases even as the size or the concentration of the object increases, and the difference is not considered significant. On the other hand, 120 kVp protocol gives the lowest detectability. The difference between the lowest and the highest detectability is very significant i.e. $p < 0.01$ for both head and body mode cases.

4.5.3 Estimation Task: Correct Estimation Ratio (CER)

For the pure contrast estimation task, 80 kVp provides the best image quality in terms of the correct estimation ratio (CER) in the head mode with varying object sizes (Fig. 4.10 (a)), and the 100 kVp is second followed by the 120 kVp. The differences of image quality among three protocols are significant. The lower the kVp protocol is, the better the image quality for the head mode case under the pure contrast estimation task. For the body mode case (Fig. 4.10 (b)), the image quality does not change significantly as the voltage lowered from 100 kVp to 80 kVp, but the difference between the highest image quality and the lowest image quality (120 kVp) is substantial.

4.5.4 Combined Task of Detection Task and Estimation Task: EAUC

In addition to pure detection and pure contrast estimation, it will be interesting to see the performance of different protocols under the combined task of detection, size and contrast estimation. For both the head and body mode, the lower the kVp, the higher the EAUC value i.e. the image quality. This result (Fig. 4.11) shows that the observer could detect the contrast-enhanced objects easier and determine the size and the contrast with higher CER under this hybrid task. The difference between the highest image quality and the lowest one is significant. This result also suggests that there is a lot of image quality improvement when the voltage lowered down from 120 kVp.

(a) (b)

Figure 4.1: (a) Head mode Gammex phantom(20 cm in diameter) (b) Body mode Gammex phantom (32 cm in diameter). The white cylinder bottles were filled with water and were inserted in the most outside of the body mode phantom.

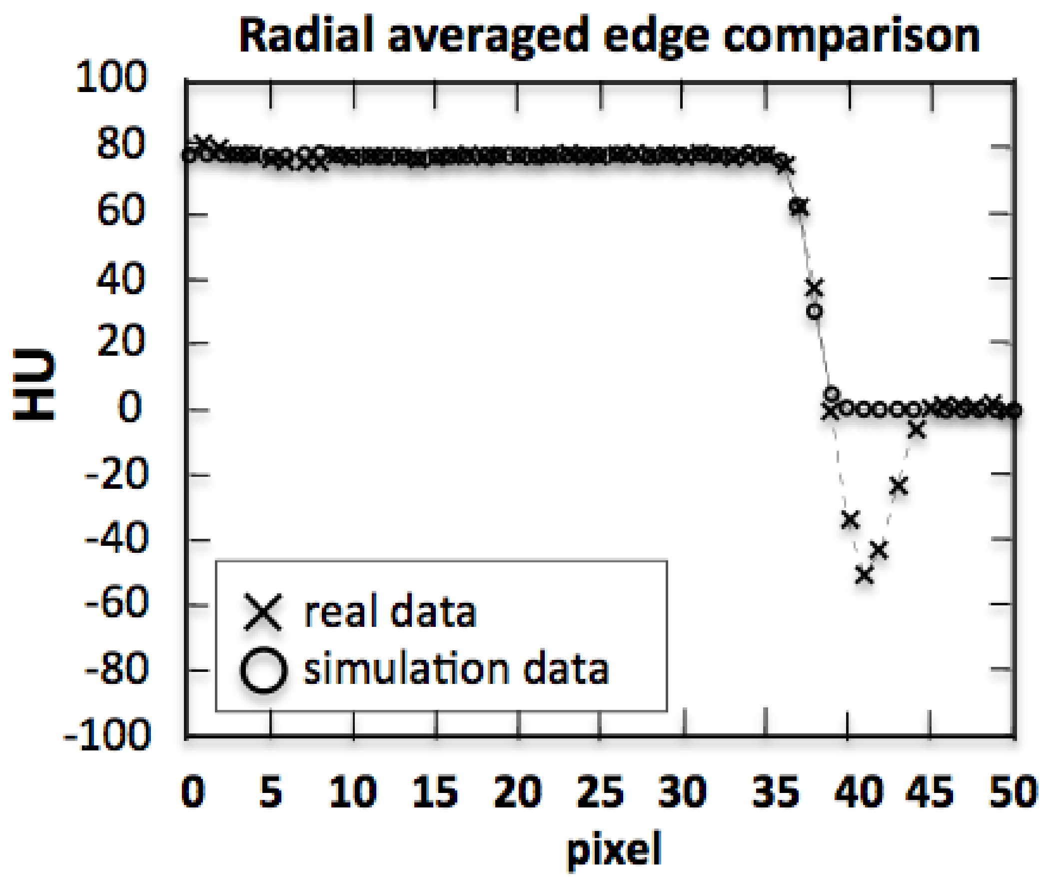

Figure 4.2: Example of comparison between edge profile of real signal and edge profile simulation signal.

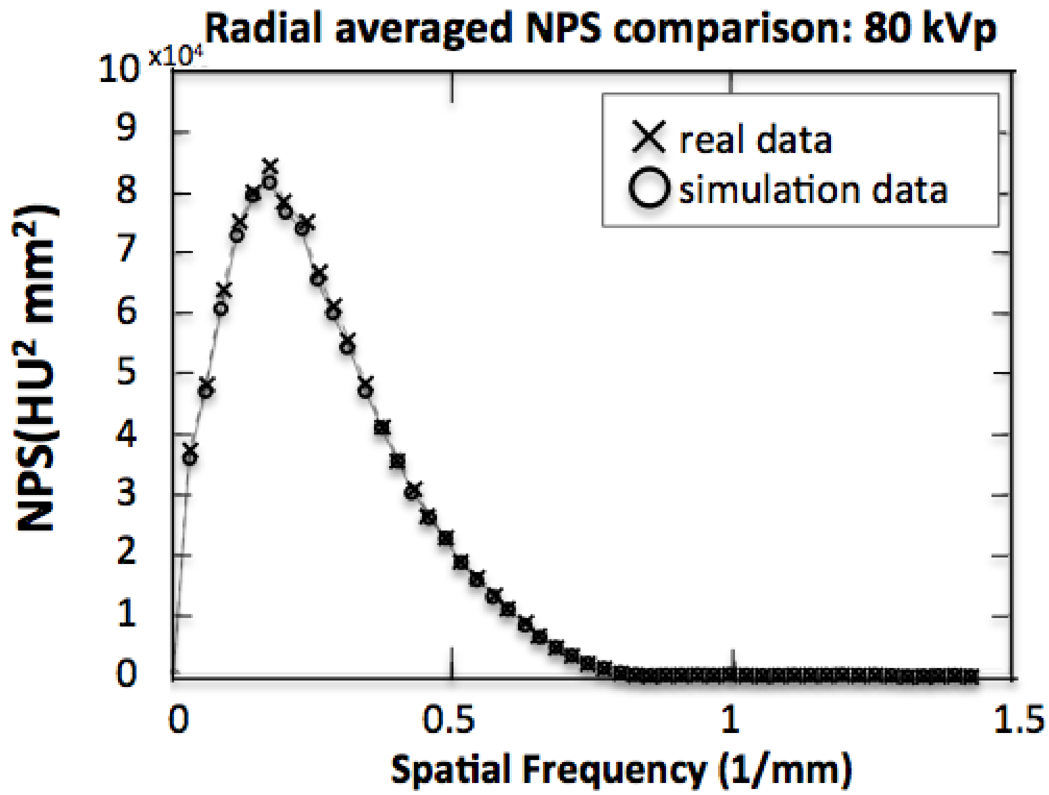

Figure 4.3: Example of comparison between NPS of real signal and NPS of simulation signal.

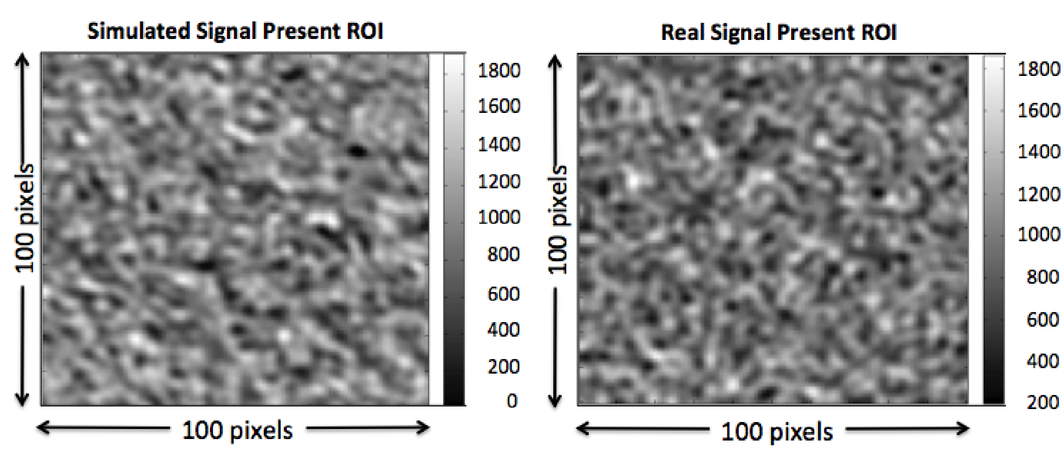

Figure 4.4: Comparison between real image and simulation image at 80kVp. Left: simulated signal absent image. Right: real signal absent image.

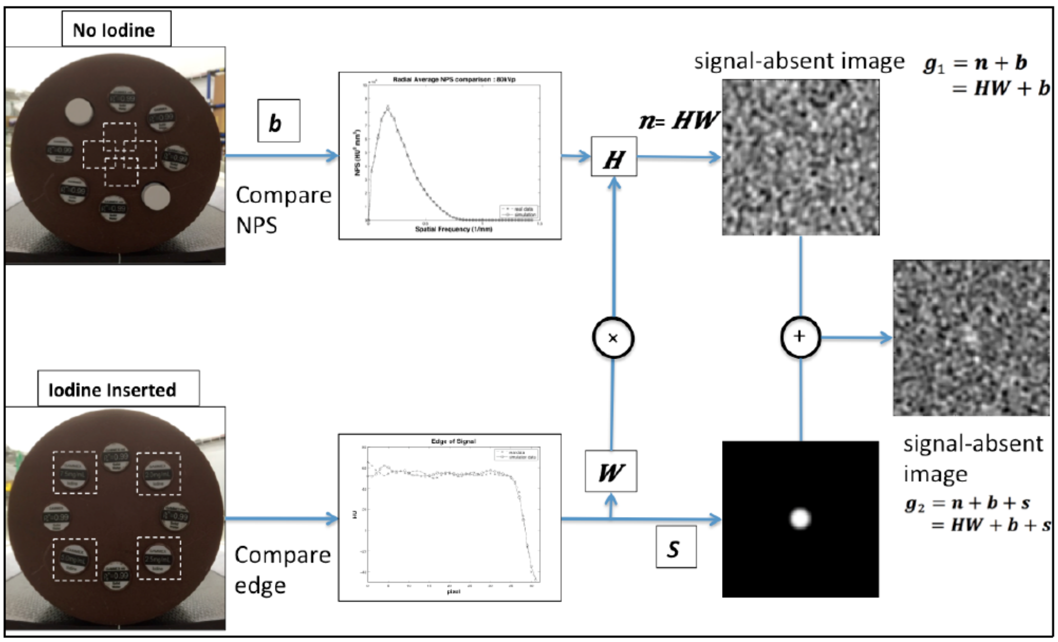

Figure 4.5: Flow chart of generation of simulation images. Phantom with only water cylinders and water bottles inserted were scanned and white dashed-line boxes represent the ROI cropped for estimate of background b, noise power spectrum (NPS) and matrix H. Phantom with iodine cylinders of different concentrations inserted were scanned and white dashed-line boxes represent the ROI cropped for simulation of signals s and vector W. Signal-absent image and signal-present image were then generated according to equation 4. 1 and 4.2.

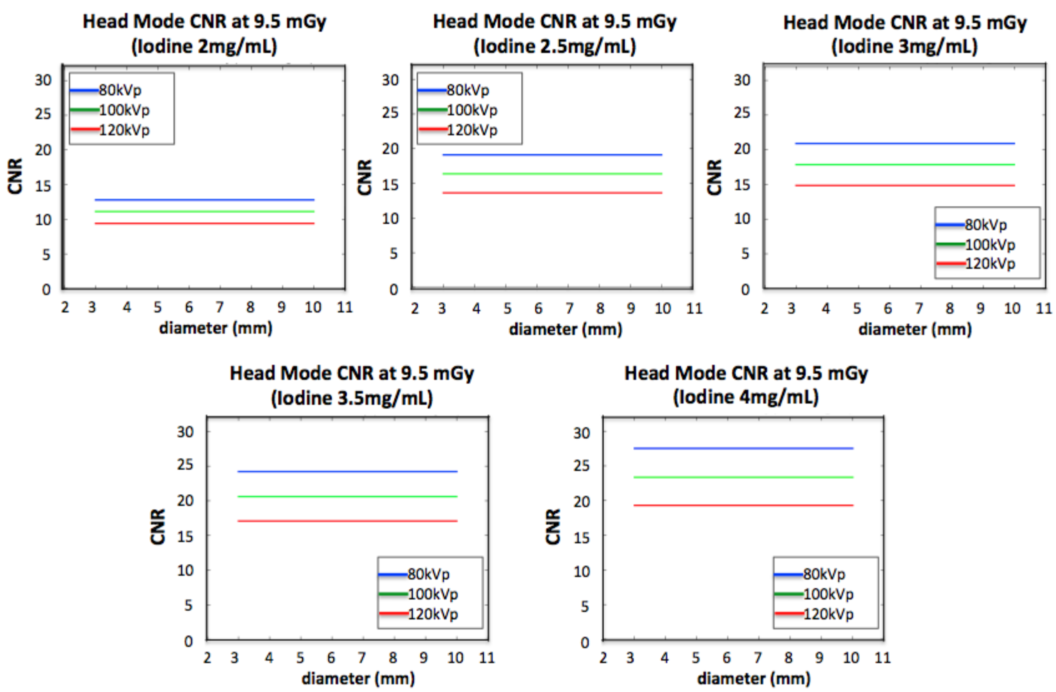

Figure 4.6: CNR value vs. the size of object (Head mode at 9.5 mGy). (a) 2mg/mL (b) 2.5 mg/mL (c) 3mg/mL (d) 3.5mg/mL (e) 4mg/mL.

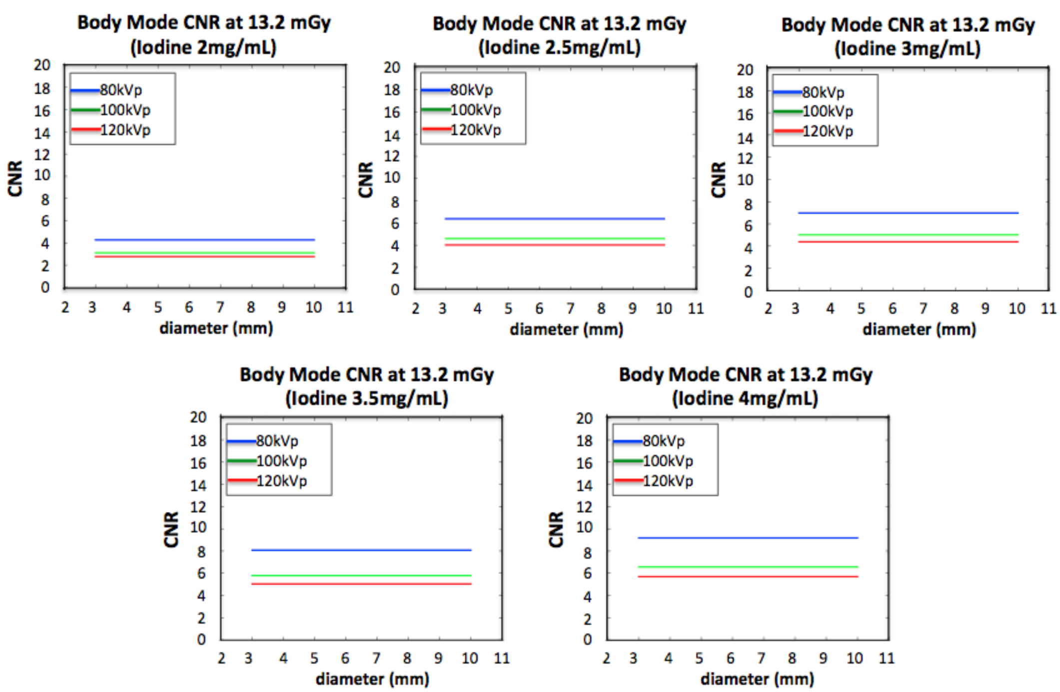

Figure 4.7: CNR value vs. the size of object (Body mode at 13.2 mGy). (a) 2mg/mL (b) 2.5 mg/mL (c) 3mg/mL (d) 3.5mg/mL (e) 4mg/mL.

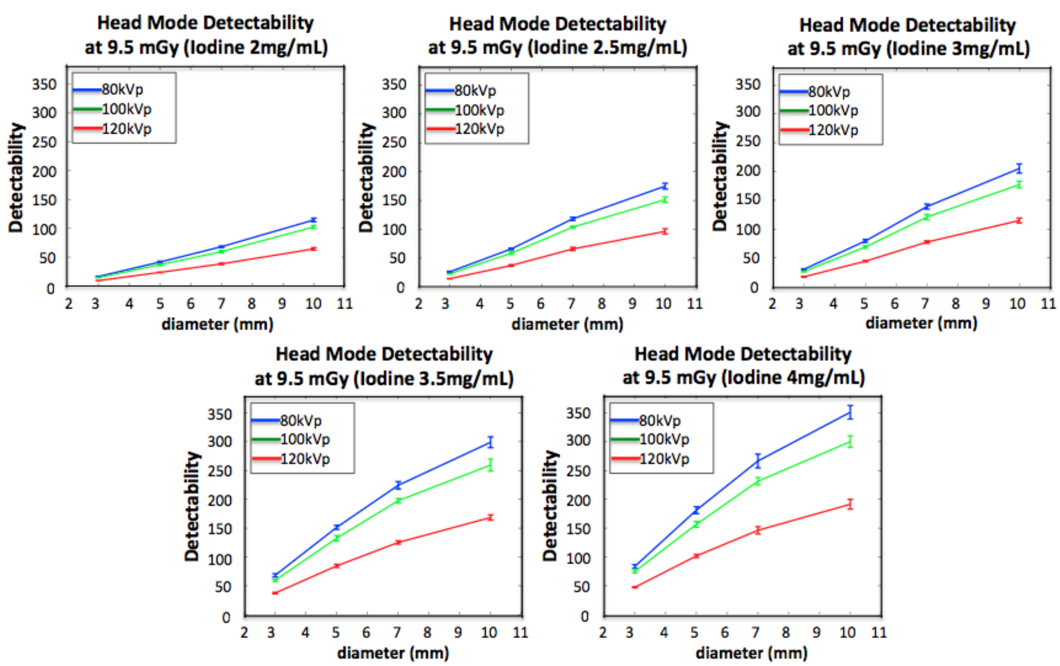

Figure 4.8: Detectability vs. the size of object (Head mode at 9.5 mGy). (a) 2mg/mL (b) 2.5 mg/mL (c) 3mg/mL (d) 3.5mg/mL (e) 4mg/mL.

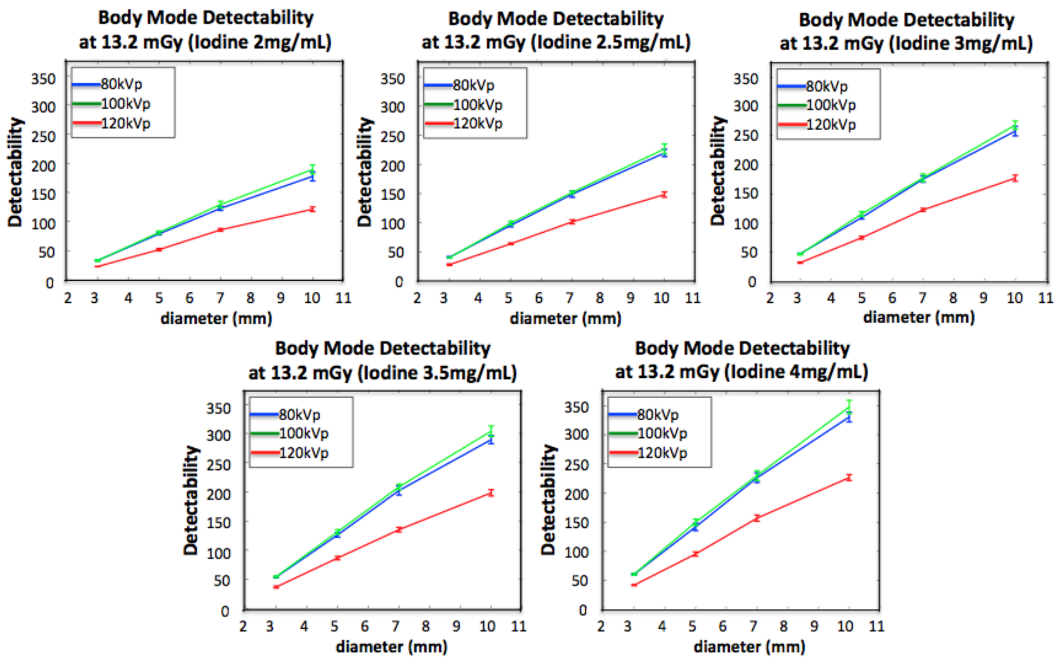

Figure 4.9: Detectability vs. the size of object (Body mode at 13.2 mGy). (a) 2mg/mL (b) 2.5 mg/mL (c) 3mg/mL (d) 3.5mg/mL (e) 4mg/mL.

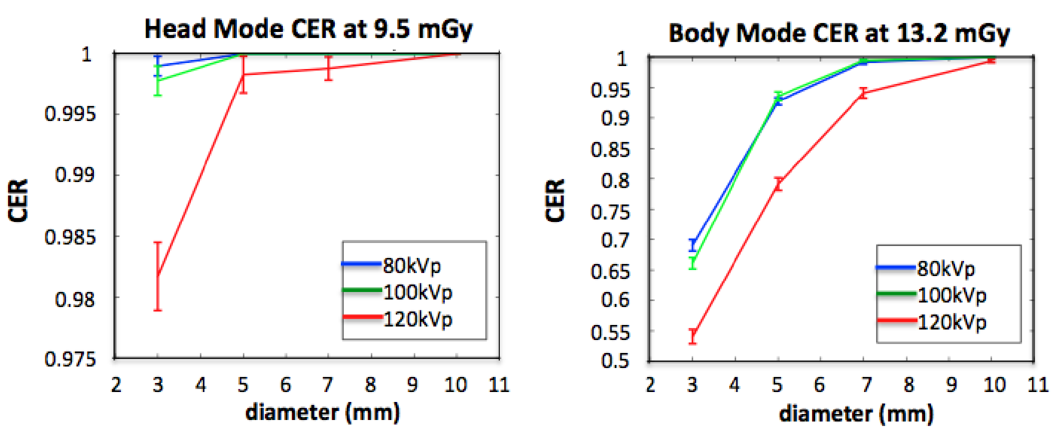

Figure 4.10: CER value vs. the size of object. (a) Head mode at 9.5 mGy. (b) Body mode at 13.2 mGy.

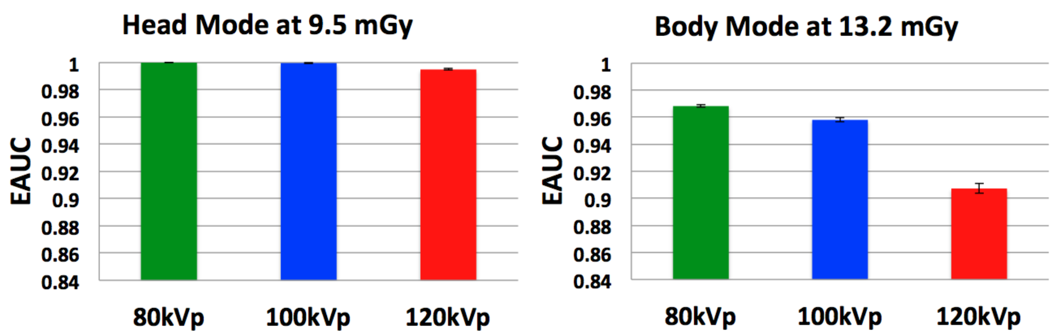

Figure 4.11: EAUC value. (a) Head mode at 9.5 mGy. (b) Body mode at 13.2 mGy.

CHAPTER 5

Quantitative Temporal Resolution Evaluation in Cardiac CT

There is a strong motivation to make improvements in cardiac imaging due to cardiovascular disease (CVD) becoming the leading source of death in western countries. Computed tomography (CT) has become a highly demanded tool in cardiac imaging—particularly for the imaging of coronary arteries (Schroeder et al. (2007)). In clinical CT, it is very challenging to image the human heart at very low dose values, at high temporal resolution, nearly free of motion artifacts, and in scan time shorter than a single breath-hold. Various approaches have been developed for dose reduction. Today, scans around or below 1 mSv have become possible in routine use. Several methods, such as the early attempt of prospective gating and faster scans from the continuous rotations provided by spiral CT techniques, have been developed for the improvement of image quality of coronary arteries as well. However, these methods were not considered encouraging and did not making a breakthrough in Cardiac CT (Shemesh et al. (1995); Boyd and Lipton (1983)). The other approach introduced in late 2005 was the Multi-source acquisitions or dual source CT (DSCT) technologies (Hsieh and Senzig (2002); Petersilka et al. (2008); Flohr et al. (2006); KachelrieB et al. (2006)), and it might be considered as the last major step for static images of imaging human hearts. However, artifacts still may exist even if the multi-source acquisitions and faster gantry rotation are applied. Research (Rohkohl et al. (2012)) has indicated that improvements could be reached through algorithmic motion estimation on dual-source CT scanners. Thus, an efficient way via algorithmic compensation was considered. There are a variety of algorithms, which can be separated into two categories, that have been designed for reducing the artifacts in heart motion images. The first type is the iterative method that uses a subset of the image data and priori knowledge (Chen et al. (2009); Schondube et al. (2012)). The second type is the motion compensation method (Cho et al. (2012);

Iatrou et al. (2010); Tang et al. (2009)) that uses the information of the desired cardiac phase and one or two closer neighboring phases. In this study, we focused on the image quality evaluation of a motion compensation algorithm developed by GE (SnapShot Freeze). SnapShot Freeze (SSF) utilizes the information from adjacent cardiac phases within a single cardiac cycle to characterize the motion of a vessel and make a proper compensation for the vessel motion at the prescribed phase in the image space. SSF images improve the temporal resolution and improve visualization of the vessel and plaque estimation. In addition to reconstruction algorithms, we also investigated the effect of phase selection before the image reconstruction process. To quantitatively measure the improvements in image quality, a clinical related task-based measurement has been developed to estimate stenosis.

5.1 Motion Correction Algorithm

The main source of artifacts in Cardiac CT comes from motion. When imaging small objects such as the coronary vessels, things become very difficult as they move during data acquisition. To have better image quality, high temporal resolution is needed in Cardiac CT scanners to freeze the motion of objects as a snapshot.

The temporal resolution is about 150 ms for the single-source CT system with the gantry rotation time 0.3 s. The speed of the coronary arteries could reach 100 mm/s or even more (Achenbach et al. (2000)). Consequently, a blurring in the same order of diameter of the vessel occurs during data acquisition. Therefore, it could cause severe mistakes in diagnosis. This problem becomes more serious when imaging hearts with a Cardiac CT scanner with a rotation time of 0.5 s.

The motion compensation software SSF developed by GE was designed towards artifacts in coronary CT angiography. It combines the two important techniques-non-rigid deformations of coronary tree (Bhaglia et al. (2012)) and splitting motion compensation into linear segments of local motion (Pack and Claus (2011)). The purpose of deformations of coronary tree is dedicated to improve the temporal resolution as well as to improve the utilization of dose by mapping all cardiac phases into one phase. SSF uses the information from adjacent cardiac phases within a single cardiac cycle to track both the path and the velocity of the vessel. This, therefore, helps the motion compensation algorithm compress the temporal window in a more efficient way. SSF could adaptively compress the temporal window in a local region where the compression is most needed. SSF also corrects each voxel of the motion vessel of different degrees of motion.

5.2 Mathematic Model Observer and Figure of Merit

In this study, a task-based image quality measurement has been developed and applied to Cardiac CT images. To evaluate the temporal resolution of image reconstruction algorithms, a threshold-based observer was employed to estimate the three different percentage plaques in a motion vessel with two different heart-beating rates - 60 and 80 bpm. There was a threshold according to a motion-free reference image and these three different percentage plaques. The method of calibrating the threshold will be illustrated in the section 5.3. The observer estimated the percentages of the plaques in the testing images according to the total number of pixels that are higher than a calibrated threshold. The estimated percentages were then compared to the true values. The figure of merit, ensemble mean square error (EMSE), was then calculated for the deviations between estimated values and true values. The lower the EMSE is, the better the image quality.

5.3 Phantoms and Data Acquisition

The phantoms used in this research have two parts: (1) motion phantom (Fig. 5.1) and (2) vessel phantom (Fig. 5.2). The motion phantom used in this study was the MOCOMO phantom developed by Fuyo Corporation, Japan and John Hopkins University, U.S.A. The designer was Dr. Armin Arbab-Zadeh MD, John Hopkins University, U.S.A. This phantom is designed for the evaluation of coronary artery depiction. The shift in z-axis (table) is replicated and contorted with pulse emission synchronized electrocardiogram (ECG). The MOCOMO phantom was filled with water in our experiments. A vessel phantom of 3 mm in diameter was hung inside the MOCOMO phantom by elastic bands. The vessel phantom was rotated relative to z- axis (table) by the MOCOMO phantom in the motion mode.

In this work, 25%, 50%, and 75% plaque imbedded in a vessel phantom of 3 mm in diameter were imaged on a 64-slice commercial CT scanner (GE HD Discovery 750). The X-Ray voltage was kept at 120 kVp for all scans. A helical scan at 600 mAs was acquired for the generation of an image template. 26 identical scans under the Cardiac scan mode at 300 mAs were acquired at three different heart beating rates (1) no heart beating (static) (2) 60 bpm and (3) 80 bpm. 12 individual realizations of the moving signal were extracted by vessel tracking from each scan. Overall, there were 312 individual images made for each plaque. The images were then reconstructed at a field of view of 200 mm with a matrix size of 512×512 image pixels. ROIs with a size of 100×100 pixels containing plaques in the center were cut out from images, which was about 39 mm 39 mm in size.

5.4 Calibrations and Estimation Process

The image template (Fig. 5.3) or the reference is a mean image averaged by a helical scan at 600 mAs without plaque but only with contract agent (Iodine). To avoid any errors of fabricating plaques, calibrations were necessary before the analysis. The calibration process was the following: (1) Count the number of pixels of the contrast agent (the white area in Fig. 5.3) in the reference image (Fig. 5.1) with an initial threshold (2) Generate the mean images of 25%, 50%, and 75% plaque averaged by 312 images extracted from a static vessel scanned under Cardiac mode at 300 mAs. Directly applying the mean images of 25%, 50%, and 75% (Fig. 5.4) plaque on the reference image. (3) Count the number of pixels above the initial threshold (CT number of contrast agent) and calculate the ensemble mean square error (EMSE) based on the difference between experimental values and ideal values (25%, 50%, and 75%). (4) Tune the threshold to reach the lowest EMSE value. The calibration results show that the threshold is about 1230 HU. The calibration indicated that

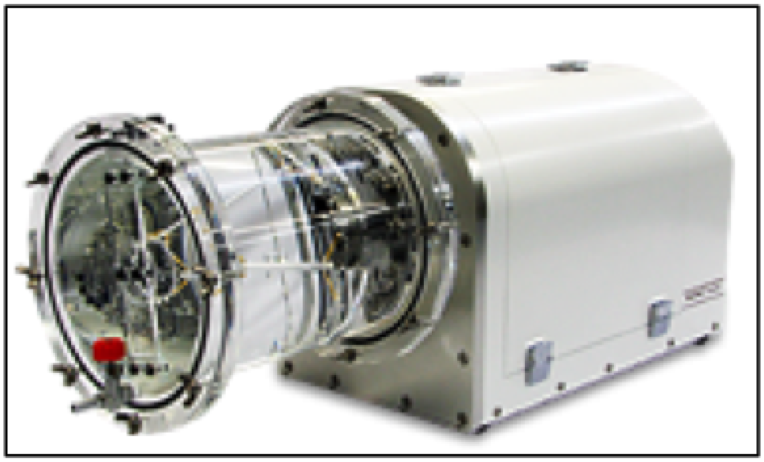

Figure 5.1: MOCOMO phantom. MOCOMO phantom is designed for the evaluation of coronary artery depiction. The shift in z-axis is replicated and contorted with pulse emission synchronized electrocardiogram (ECG).

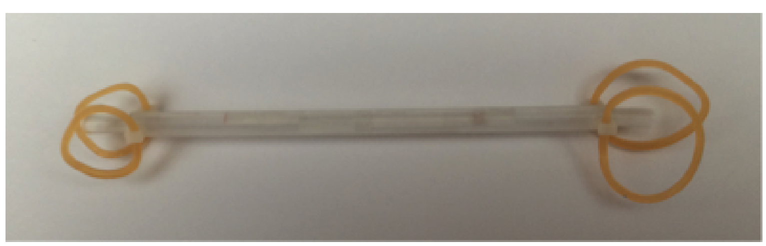

Figure 5.2: Vessel phantom 3 mm in diameter. There are four elastic bands (yellow) for hanging inside the MOCOMO phantom.

instead of 25%, 50%, and 75% stenosis, the percentage of stenosis is about 24.7997% (Fig. 4 Left), 52.68% (Fig. 5.4 Middle), and 89.3029% (Fig. 4 right) respectively. So from now on, we will use 24.80%, 52.68%, and 89.30% instead of 25%, 50%, and 75% to be our baseline/true value for the entire analysis.

For testing, images from a motion vessel at both 60 bpm and 80 bpm acquired at 300 mAs were scanned by the Cardiac mode. Two different reconstruction phases (cardiac cycle) were used before image reconstructions: (1) the default phase and (2) the superior phase. The superior phase is selected empirically. Images were generated by two algorithms:(1) FBP and (2) SSF for all phases and heart-beating rates. Two heart-beating rates, two reconstruction phases, and two image reconstruction algorithms resulted in 8 kinds of testing images for our observer study. Each category has 312 testing images. Each image was compared to the reference, and the number of pixels above the calibrated threshold (1230 HU) was counted. Estimated plaque percentage was based on the total number of pixels counted. The EMSE value of each category then was calculated based on how much deviation of the estimate value from the calibrated baseline/true value was. The lower the EMSE value is, the lower the estimate error i.e. the better the image quality.

5.5 Measurements and Analyses

There were three different percentages of stenosis under the estimation task by the mathematic observer. Two heart-beating rates (60 bpm and 80 bpm), two reconstruction phases (default and superior), and two image reconstruction algorithms (FBP and SSF) were given in our observer study. Hence there were 8 classes total for the temporal resolution evaluation. The observer estimated the percentages of plaque according to the calibrated thresholds for each class. To compute the estimate, each class used 312 testing images. Dots and lines in Fig. 5.5 and Fig. 5.6 represented the estimate values and baseline values, respectively. Among cases of 60 bpm, all of them (Fig. 5.5(a) (d)) indicate that the accuracy of estimating 52.68% plaque is higher than the other two (24.80% and 89.30%). The performance

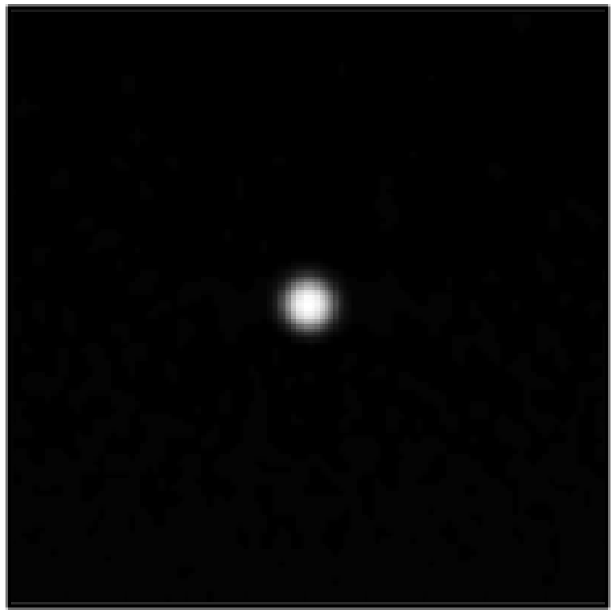

Figure 5.3: Image template used as a reference for threshold-based observer study. The image template was averaged by images from a helical scan acquired at 600 mAs.

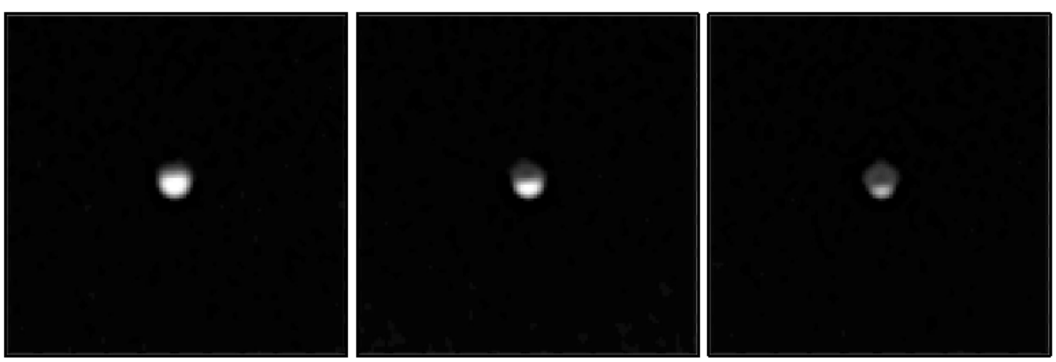

Figure 5.4: Image used for calibration. Images were averaged by 312 images (each case) extracted from a static vessel scanned under Cardiac mode at 300 mAs. Left: 25% plaque. Middle: 50% plaque. Right: 75% plaque.

of 60 bpm FBP (Fig. 5.5(a)) at the default phase is very similar to the performance of 60 bpm SSF at the default phase (Fig. 5.5(c)). Similar phenomenon could be observed as well on Fig. 5.5(b) and Fig. 5.5(d). This implies that as selected at the same reconstruction phase, FBP and SSF image at 60 bpm have very similar performance. In other words, when the same algorithm is used, the estimation on the 52.68% plaque is more accurate when using images selected from the superior phases case (Fig. 5.5 (b) and (d)) than images selected from the default phase case (Fig. 5.5 (a) and (c)). The better estimation on 52.6786% plaque nearly explains why the quality of the superior phase images is better than the quality of the default phase images.

From the Fig. 5.6, the 80 bpm images reconstructed by FBP selected at the default (Fig. 5.6 (a)) and the superior (Fig. 5.6 (b)) reconstruction phase both show that several estimated values deviated far away the baseline values. In comparison, when using the SSF algorithm on the same images, a decrease in difference between the estimated values and the baseline values was observed, for both the default and superior cases. Similar, for the 60 bpm case, when comparing the images generated by the same reconstruction algorithm but selected at different reconstruction phases, the cases of superior phase provide better performance on estimating 52.68% plaque.

In order to quantify the image quality improvements, the EMSE values for all 8 classes have been calculated and plotted in the bar charts (Fig. 5.7). The lower the value is, the better the image quality. There are significant improvements in image quality when the cardiac cycle phase was selected superiorly and images were reconstructed by the motion compensation algorithm SSF.

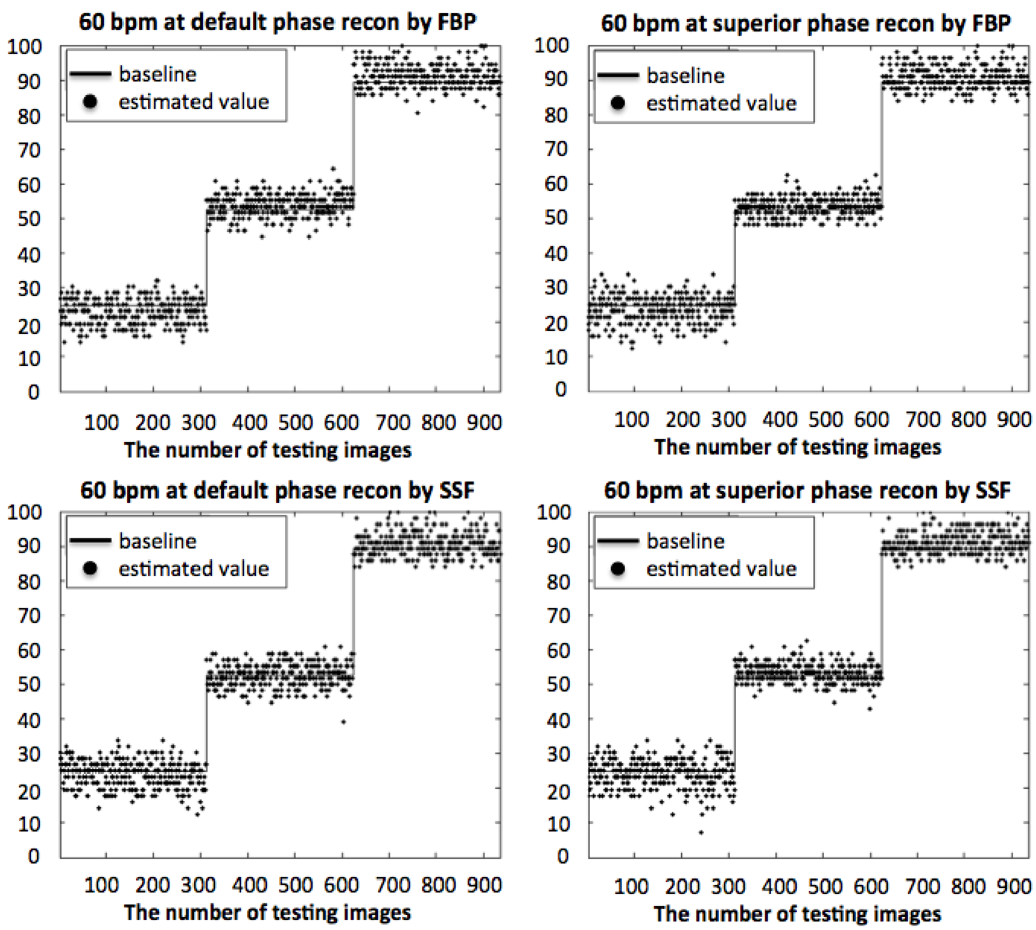

Figure 5.5: Scattering plots of 60 bpm case. Solid line refers to the baseline and dot stands for estimate value. (a) Default phase with FBP (b) Superior phase with FBP (c) Default phase with SSF (d) Superior phase with SSF.

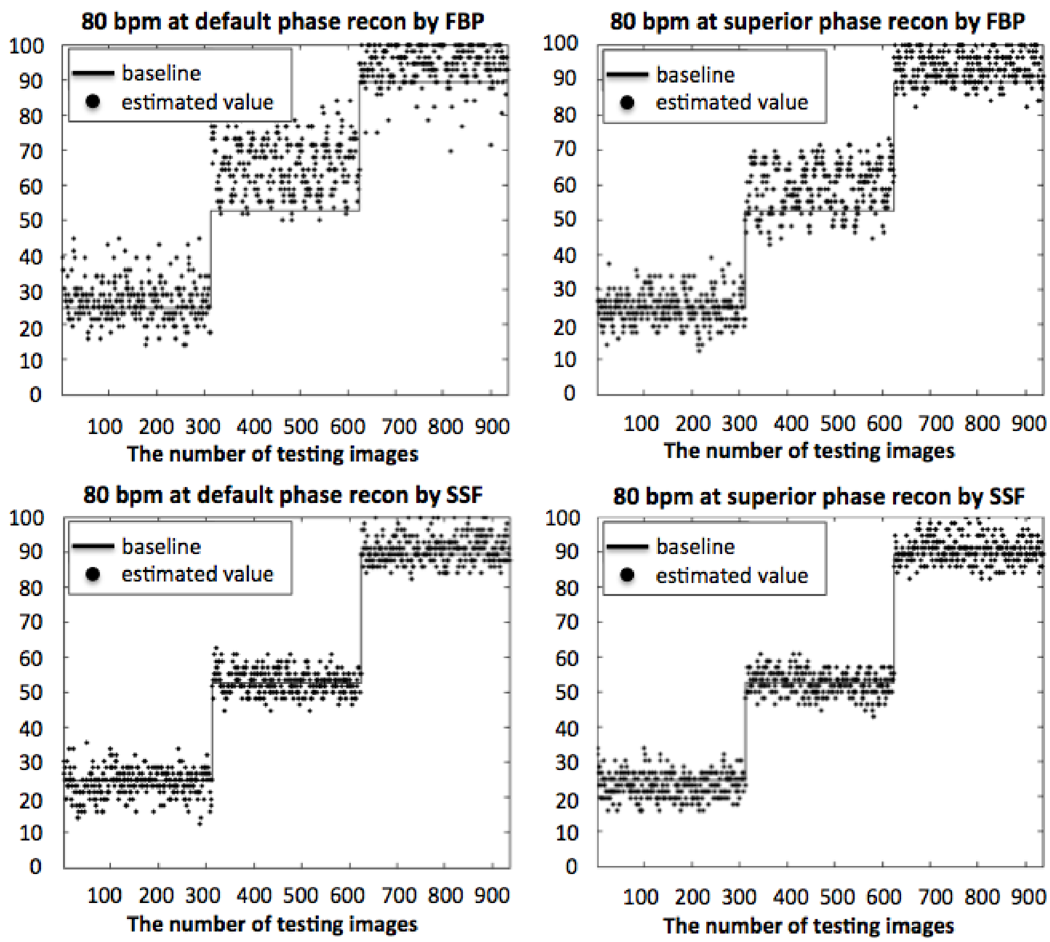

Figure 5.6: Scattering plots of 80 bpm case. Solid line refers to the baseline and dot stands for estimate value. (a) Default phase with FBP (b) Superior phase with FBP (c) Default phase with SSF (d) Superior phase with SSF.

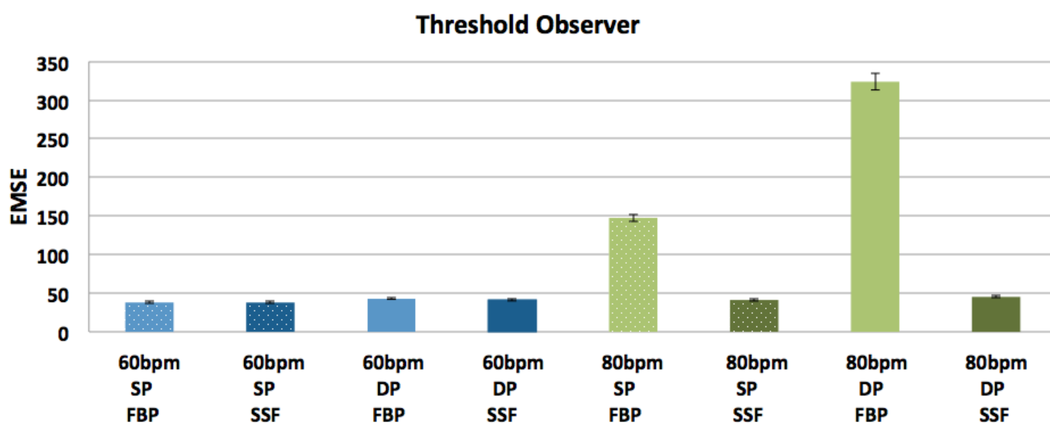

Figure 5.7: EMSE values and associated error bars for all cases. (SP=Superior Phase; DP=Default Phase)

CHAPTER 6

Use of Task-based Assessment in CT Quality Control

For detection tasks and combined detection/estimation tasks in which the observer must detect a signal and estimate parameters of the signal (e.g., contrast, size), the channelized Hotelling observer (CHO) and channelized scanning linear observer (CSLO) have been successfully utilized to measure image quality. To implement these two observers, first-order and second-order statistics are necessary, which are typically estimated using sample images from a training dataset. The goal of this chapter is to explore different methods to reduce the data demands for computing CHO/CSLO performance and propose a practical model observer based approach that can be used in the clinical environment for routine image quality assessment of CT imaging systems. We will use high dose images to estimate the first-order statistics, i.e., the mean signal image, and use the Leave-One-Out Covariance (LOOC) (Hoffbeck and Landgrebe (1996)) method to estimate the second-order statistics, i.e., the covariance matrices.

6.1 Real Data Preparation

The details of phantoms have been described in the previous chapters so the information of each phantom will be skipped here. The MITA (CCT 183) phantom was used for pure detection tasks and the validation of our proposed ideas. The QRM-LC-FD1 was used for combined detection and size estimation tasks, the QRM-LC-FD4 was used for combined detection and contrast estimation tasks, and the MTIA (CCT 189) phantom was used for combination of detection, size and contrast estimation tasks. The properties of LC objects in these four phantoms have been discussed in chapter 2 and 3. Only axial scans were used in this work. The slice thickness was 0.625 mm and the collimator aperture used was 20 mm. The X-Ray voltage was kept at 120 kVp and current range was tuned to ensure that the task being performed was neither too difficult nor too easy.

Various radiation dose protocols were used and for each dose level, 50 repeat scans were acquired. The image reconstruction algorithm used in this study was filtered backprojection (FBP). From each of these scans, 10 individual realizations of signal and noise regions of interest (ROI) were extracted from each scan for each different low-contrast signal. Thus, there were 500 individual image pairs for each LC object at every dose level. The images were reconstructed at a field of view of 180 mm with a matrix size of 512 × 512 image pixels. ROIs with a size of 100 × 100 pixels containing the LC object under study in the center (signal present case) and the noise (signal absent case) were extracted from the image slices, which is about 35mm × 35mm in size.

6.2 Simulated Data Preparation, Analytical Solution, and Baseline

Since the number of acquired images is always limited in experiments, simulation data were generated to find the baseline performance and help understand the variability in our methods. Because the two key components, first- and second-order statistics, needed in our observer models are the same for pure detection tasks and combined detection and estimation tasks, without loss of generality, our simulation were focused on the case of pure detection task. The methods of generating simulated image pairs for signal-absent and signal-present classes have been addressed in chapter 4 and will not be described here again. The AUC can be approximated through the SNR of the test statistic using,

$$SNR_\lambda = \sqrt{\boldsymbol{s}^t \boldsymbol{K}_n^{-1} \boldsymbol{s}} = \sqrt{\boldsymbol{s}^t (\boldsymbol{R}_g - \boldsymbol{b}\boldsymbol{b}^t)^{-1} \boldsymbol{s}} \tag{6.1}$$

$$AUC = \frac{1}{2}(1 + \mathrm{Erf}(\frac{SNR_\lambda}{2})) \tag{6.2}$$

where SNR is the signal to noise ratio, and Erf is the error function. Since $\boldsymbol{R}_g$ is a function of dose level so the SNR and AUC are also functions of dose level.

To validate the simulated images, the radial average noise power spectrum (NPS) (Dobbin et al. (2006)) profile of the simulation and the real images were compared. The comparisons (Fig. 6.1) indicate that the simulated noise background is very similar to the real noise background. An example of a simulated image and a real image is shown in figure 6. 2. The analytically derived AUC values (solid line) were compared with the AUC values calculated using the CHO method from 20000 pairs of simulation images (dashed line)(Fig. 6. 3), and the comparisons demonstrate that our derivation of the analytical solution is reasonable.

To better understand any biases present using the real data, we compared the AUC values generated using finite numbers of simulated images with the AUC values generated by the analytical solution and 20000 pairs of simulated images. We found that the AUC values of the CHO using 240 training images and 260 testing images (solid line with triangle) were very close to both the analytical solution (solid line) and the 20000 image pairs CHO (dashed line) for all four objects with accuracy ranging from 98.25% to 99.87% (Fig. 6. 3). This high accuracy in the simulation study indicates that 500 real image pairs will result in negligible bias and can be used as the baseline in our experimental study.

6.3 Traditional Approach

The CHO and CSLO approach requires a relatively large amount of training data to estimate the means and covariance matrices to generate accurate estimates of task performance. Although the ideas and equations of CHO and CSLO have been described in the early chapters, to compare with the proposed approaches easily, we will give the equations here again.The image template calculated by the CHO is given by

$$\boldsymbol{w}_{CHO} = \boldsymbol{S}^{-1}\boldsymbol{s} \tag{6.3}$$

where $\boldsymbol{s}$ is the sample mean estimated signal and $\boldsymbol{S}$ is the common covariance matrix defined by averaging the sample covariance matrices for the two classes.

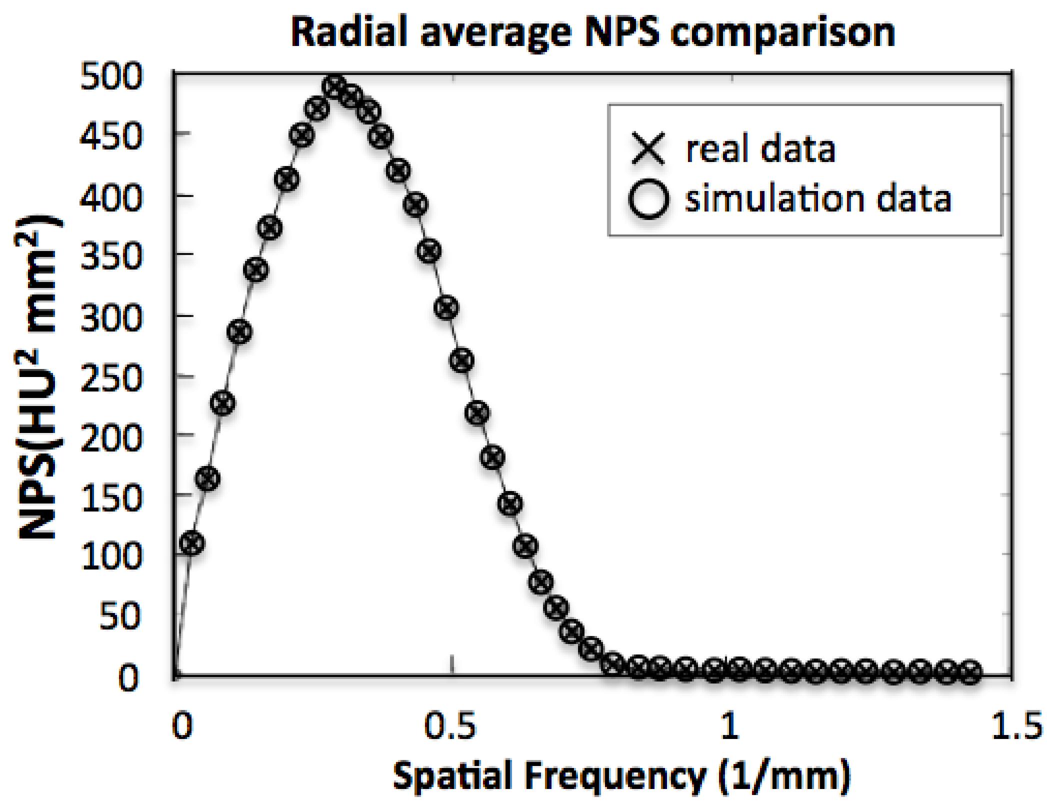

Figure 6.1: Radial average NPS profile at 120 kVp 80 mAs: (Black circle solid line) simulation data (Black dashed cross line) real data.

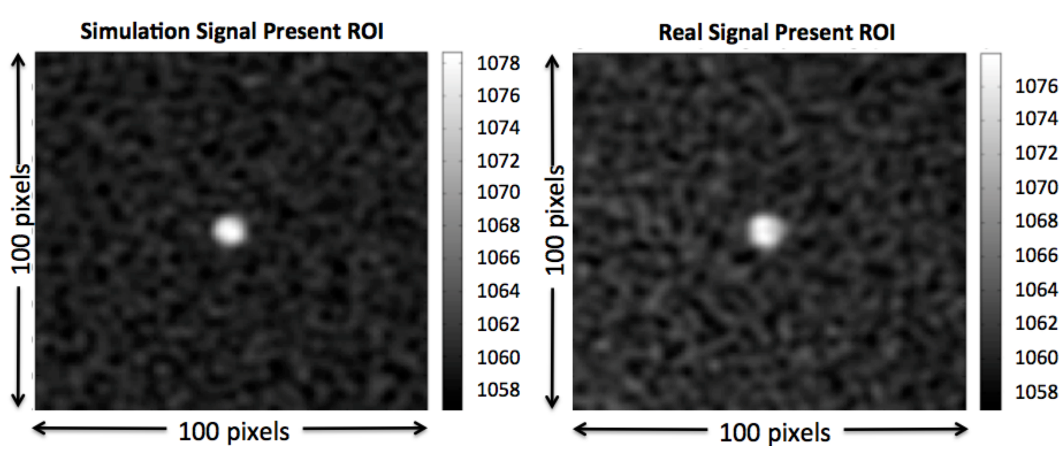

Figure 6.2: Signal-present image (3 mm 14 HU) averaged by 500 images at 120 kVp 80 mAs (Left) Simulated image (Right) Real image.

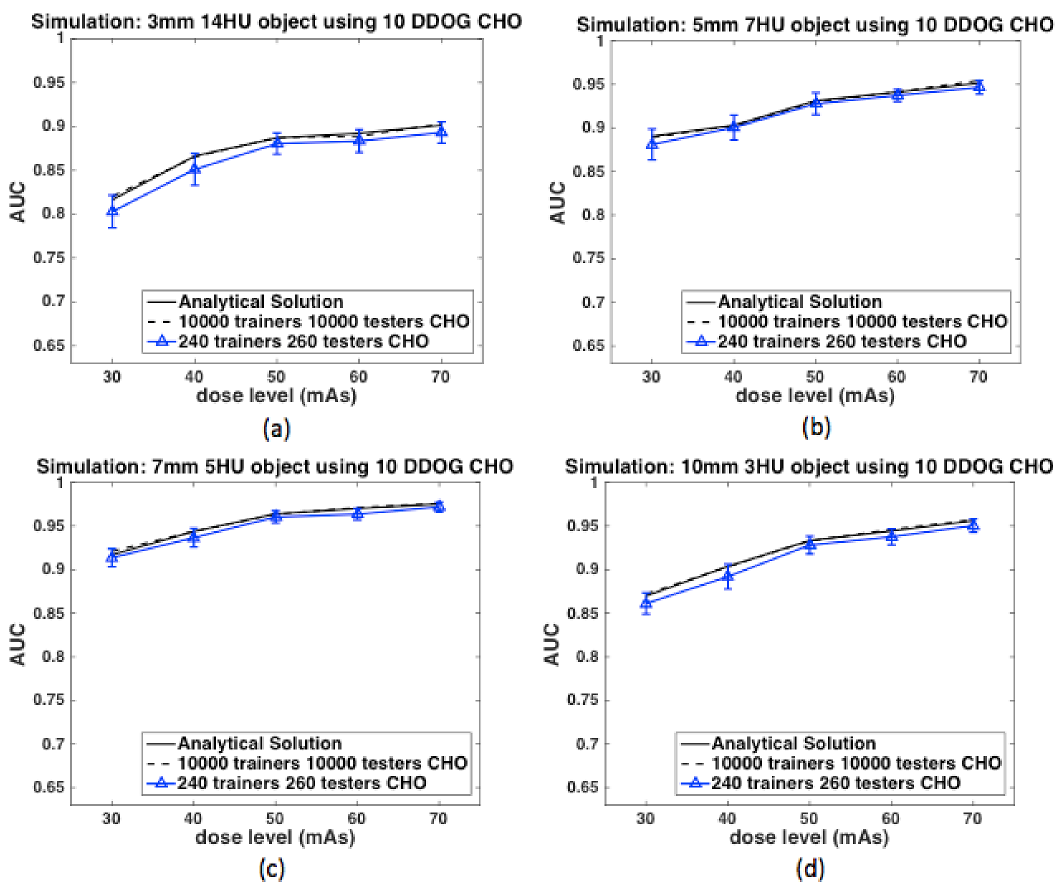

Figure 6.3: The comparison of AUC values of analytical solutions, 20000 pairs of image CHO and 500 pairs of image CHO. (a) The object of 3 mm 14 HU (b) The object of 5 mm 7 HU (c) The object of 7 mm 5 HU (d) The object of 10 mm 3 HU.

They are computed by the following equations.

$$\boldsymbol{s} = \boldsymbol{m}_2 - \boldsymbol{m}_1 \tag{6.4}$$

$$\boldsymbol{m}_i = \frac{1}{N_i}\sum_{k=1}^{N_i} \boldsymbol{x}_{i,k} \tag{6.5}$$

$$\boldsymbol{\Sigma}_i = \frac{1}{N_i - 1}\sum_{k=1}^{N_i} (\boldsymbol{x}_{i,k} - \boldsymbol{m}_i)(\boldsymbol{x}_{i,k} - \boldsymbol{m}_i)^t \tag{6.6}$$

$$\boldsymbol{S} = \frac{1}{L}\sum_{i=1}^{L} \boldsymbol{\Sigma}_i \tag{6.7}$$

where the subscript index i=2,1 refers to signal–present and signal–absent classes, respectively, and $\boldsymbol{m}_i$ is the sampling mean of channelized data set $\boldsymbol{x}_i$ for ith class, $\boldsymbol{\Sigma}_i$ is the sampling covariance matrix for the ith class, N_i is the total number of samples for ith class, and L is the number of classes. Since there are only two classes so L=2 in the pure detection task.

As with the CHO, the first- and second- order statistics need to be calculated before generating the figure of merit for CSLO. The concept of the scanning linear estimate have been provided in detail in chapter 3. The final equation of CSLO in the traditional approach can be expressed as

$$\hat{\boldsymbol{\theta}}_{CSLO}(\boldsymbol{x}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \left\{ \boldsymbol{s}^t(\boldsymbol{\theta})\boldsymbol{S}^{-1}\boldsymbol{x}^{tester} - \frac{1}{2}\boldsymbol{s}^t(\boldsymbol{\theta})\boldsymbol{S}^{-1}\boldsymbol{s}(\boldsymbol{\theta}) \right\} \tag{6.8}$$

6.4 Proposed Approach

In this section we will develop methods for computing observer performance that rely on less real data and can thus be computed more readily. Specifically we will study whether we can use 40 images from one high–dose acquisition to estimate the signal and 20 images at the appropriate dose to estimate the covariance matrix S. In the following sub-sections, the methods we used to estimate signal and covariance matrix will be illustrated.

6.4.1 Estimated Signal

Unlike the traditional method, the proposed estimated signal is given by

$$\boldsymbol{s}_{\text{High-Dose}} = (\boldsymbol{m}_{2,\text{High-Dose}} - \boldsymbol{m}_{1,\text{High-Dose}}) \tag{6.9}$$

where $\boldsymbol{m}_{2High-Dose}$ and $\boldsymbol{m}_{1High-Dose}$ are the mean of the channelized image data sets of signal present and signal absent cases from high dose images respectively. The calculation of $\boldsymbol{s}_{High-Dose}$ is the same as eq.(6.5) but requires less data. As with the CHO, the estimated signal with fewer samples in CSLO is given by

$$\boldsymbol{s}_{\text{High-Dose}}(\boldsymbol{\theta}) = \boldsymbol{m}_{2,\text{High-Dose}}(\boldsymbol{\theta}) \tag{6.10}$$

6.4.2 Estimated Covariance Matrix

The covariance matrix was estimated by the Leave-One-Out Covariance matrix (LOOC) method proposed by J. P. Hoffbeck and D. A. Landgrebe (Hoffbeck and Landgrebe (1996)). In pattern recognition, an important problem is the effect of limited training samples on classification performance because training data is always not abundant and expensive. When the ratio of the number of training samples to the dimensionality of data becomes small, deterioration of classification happens because the parameter estimation becomes unstable. The LOOC method developed by J.P. Hoffbeck and D. A. Landgrebe can estimate the covariance matrix when the training data is limited. The covariance matrix was estimated by the combination of a sample covariance matrix, a common covariance matrix, and the diagonal matrices associated with them. The equation of ith covariance matrix proposed by Hoffbeck and Landgrebe is given by

$$\boldsymbol{C}_i(\alpha_i) = \alpha_{i,1}\text{diag}(\boldsymbol{\Sigma}_i) + \alpha_{i,2}\boldsymbol{\Sigma}_i + \alpha_{i,3}\boldsymbol{S} + \alpha_{i,4}\text{diag}(\boldsymbol{S}) \tag{6.11}$$

where the mixing parameters $\alpha_i = [\alpha_{i,1}\alpha_{i,2}\alpha_{i,3}\alpha_{i,4}]$ are required to sum to unity

$$\sum_{j=1}^{4} \alpha_{i,j} = 1 \tag{6.12}$$

The values of the mixing parameters α_i are selected to maximize the average log Gaussian likelihood of leave-one-out samples, which is called leave-one-out likelihood (LOOL). The LOOL is computed as follows,

$$\mathrm{LOOL}_i(\alpha_i) = \frac{1}{N_i}\sum_{k=1}^{N_i} \ln\Big[f(\boldsymbol{x}_{i,k} \mid \boldsymbol{m}_{\mathrm{i/k}}, \boldsymbol{C}_{\mathrm{i/k}}(\alpha_i))\Big] \tag{6.13}$$

where subscript index i/k means the quantity is computed without using the kth samples from class i. The function f is the Gaussian likelihood

$$f(\boldsymbol{x}_{i,k} \mid \boldsymbol{m}_{\mathrm{i/k}}, \boldsymbol{C}_{\mathrm{i/k}}(\alpha_i)) = \frac{1}{\sqrt{(2\pi)^P \mid \boldsymbol{C}_{\mathrm{i/k}} \mid}} \exp\Big[\frac{-1}{2}(\boldsymbol{x}_{i,k} - \boldsymbol{m}_{\mathrm{i/k}})^t \boldsymbol{C}^{-1}_{i/k}(\boldsymbol{x}_{i,k} - \boldsymbol{m}_{\mathrm{i/k}})\Big] \tag{6.14}$$

where P is the number of pixels. In this study, P is equal to the number of channels, 10, because all the image vectors were channelized by 10 DDOG channels. The associated leave-one-out quantities are computed as follows:

$$\boldsymbol{m}_{\mathrm{i/k}} = \frac{1}{N_i - 1}\sum_{j=1, j\neq k}^{N_i} \boldsymbol{x}_{i,j} \tag{6.15}$$

$$\boldsymbol{\Sigma}_{\mathrm{i/k}} = \frac{1}{N_i - 2}\sum_{j=1, j\neq k}^{N_i} (\boldsymbol{x}_{i,j} - \boldsymbol{m}_{\mathrm{i/k}})(\boldsymbol{x}_{i,j} - \boldsymbol{m}_{\mathrm{i/k}})^t \tag{6.16}$$

$$\boldsymbol{S}_{\mathrm{i/k}} = \Big(\frac{1}{L}\sum_{j=1, j\neq i}^{L} \boldsymbol{\Sigma}_j\Big) + \frac{1}{L}\boldsymbol{\Sigma}_{\mathrm{i/k}} \tag{6.17}$$

$$\boldsymbol{C}_{\mathrm{i/k}}(\alpha_i) = \alpha_{i,1}\mathrm{diag}(\boldsymbol{\Sigma}_{\mathrm{i/k}}) + \alpha_{i,2}\boldsymbol{\Sigma}_{\mathrm{i/k}} + \alpha_{i,3}\boldsymbol{S}_{\mathrm{i/k}} + \alpha_{i,4}\mathrm{diag}(\boldsymbol{S}_{\mathrm{i/k}}) \tag{6.18}$$

Once the LOOL $\boldsymbol{C}_i$ of each class is determined, the final covariance matrix used in our approach can be calculated by the average of all covariance matrices from all classes. The final covariance matrix can be written as

$$\boldsymbol{C} = \frac{1}{L}\sum_{i=1}^{L} \boldsymbol{C}_i \tag{6.19}$$

6.5 Figure of Merit (FOM) and Relevant Mean and Variance

The input to our observer models is channelized image $\boldsymbol{x}_{tester}$. Note here that $\boldsymbol{x}_{tester}$ could be either signal present or absent and if there is a signal present, the size and contrast may vary depending upon which phantom is used. The test statistic can be described as

$$\boldsymbol{t}_l(\boldsymbol{\theta}) = \boldsymbol{w}^t(\boldsymbol{\theta})\boldsymbol{x}_l \tag{6.20}$$

$$\boldsymbol{t}_l(\hat{\boldsymbol{\theta}}_{\text{CSLO(HL-CSLO)},l}) \tag{6.21}$$

where $\boldsymbol{\theta}$ is the physical property of signal. $\hat{\boldsymbol{\theta}}_{\text{CSLO(HL-CSLO)},l})$ is given by

$$\hat{\boldsymbol{\theta}}_{\text{CSLO(HL-CSLO)},l}(\boldsymbol{x}_{tester,l}) = \underset{\boldsymbol{\theta}}{\operatorname{argmax}} \left\{ \boldsymbol{\zeta}(\boldsymbol{\theta})^t \boldsymbol{K}^{-1} \boldsymbol{x}_{tester,l} - \frac{1}{2} \boldsymbol{\zeta}(\boldsymbol{\theta})^t \boldsymbol{K}^{-1} \boldsymbol{\zeta}(\boldsymbol{\theta}) \right\} \tag{6.22}$$

where the channelized signal $\boldsymbol{\zeta}$ and the channelized covariance $\boldsymbol{K}$ can be estimated by the traditional method and proposed method. They can be described as

$$\boldsymbol{\zeta} = \begin{cases} \boldsymbol{s}_{\text{Full}}, & \text{using 240 trainers} \\ \boldsymbol{s}_{\text{Less}}, & \text{using 40 trainers} \\ \boldsymbol{s}_{\text{High-Dose, Less}}, & \text{using 40 trainers} \end{cases} \tag{6.23}$$

$$\boldsymbol{K} = \begin{cases} \boldsymbol{S}_{Full} & \text{using 240 trainers} \\ \boldsymbol{S}_{Less} & \text{using 20 trainers} \\ \boldsymbol{C}_{Less} & \text{using 20 trainers} \end{cases} \tag{6.24}$$

There are 9 possible combinations of $\boldsymbol{\zeta}$ and $\boldsymbol{K}$. Among of them, (s_{Full},S_{Full}) case is the baseline as mentioned before. Excluding the baseline, among the rest of 8 possible combinations, the ($\boldsymbol{s}_{Less}$,$\boldsymbol{S}_{Less}$) case (traditional approach with less data) and ($\boldsymbol{s}_{High-dose,Less}$,$\boldsymbol{C}_{Less}$) case (proposed approach with less data) represent the lower and upper limit respectively. To investigate the improvement of our proposed method with less data, we focused on the comparisons among baseline, ($\boldsymbol{s}_{Less}$,$\boldsymbol{S}_{Less}$), and ($s_{High-dose,less}$,$\boldsymbol{C}_{Less}$) and the other 6 combinations were ignored for all tasks.

ROC curve and EROC curve (Clarkson (2007)) can be generated after generating the test distributions of $\boldsymbol{t}_l$. The figure of merit (FOM) of CHO and CSLO is the area under ROC curve (AUC) and the area under EROC (EAUC) curve respectively. The mean values and variances of the FOM for all tasks were estimated by the shuffle method (Fukunaga and Hayes (1989)).

6.6 Validation Using Simulation Data

To verify our proposed method, three approaches were studied in both simulation and a real–data study: (1) baseline-traditional approach using 240 training samples and 260 testing samples (2) traditional approach using 20 training samples and 80 testing samples (3) proposed approach using 20 training samples and 80 testing samples. Again, first- and second-order statistics, needed to be estimated are the same for the pure detection task and the combined detection and estimation task. Without loss of generality, the validation will be focused on the case of the pure detection task using simulated data. The MITA CCT183 phantom was used for this verification study. To evaluate the performance of different approaches, the accuracy defined by the ratio of FOM calculated by the testing approach to the FOM from the baseline was used. The accuracy is in the range of 88%∼98% for the traditional method using fewer data in simulation. For our proposed method using fewer data, the accuracy is in the range of 97%∼100%. The comparisons have been shown on figure 6.4. The simulation suggests that our proposed method provides more accurate results than the traditional approach while the amount of data was reduced to 20% of the original amount of data.

6.7 Results

In our study, four tasks were investigated: (1) detection tasks (2) detection and size estimation tasks (3) detection and contrast estimation tasks, and (4) detection, size and contrast estimation tasks. For all tasks in this study, the location of signal was known exactly and was always in the center of ROI. For each task, three methods

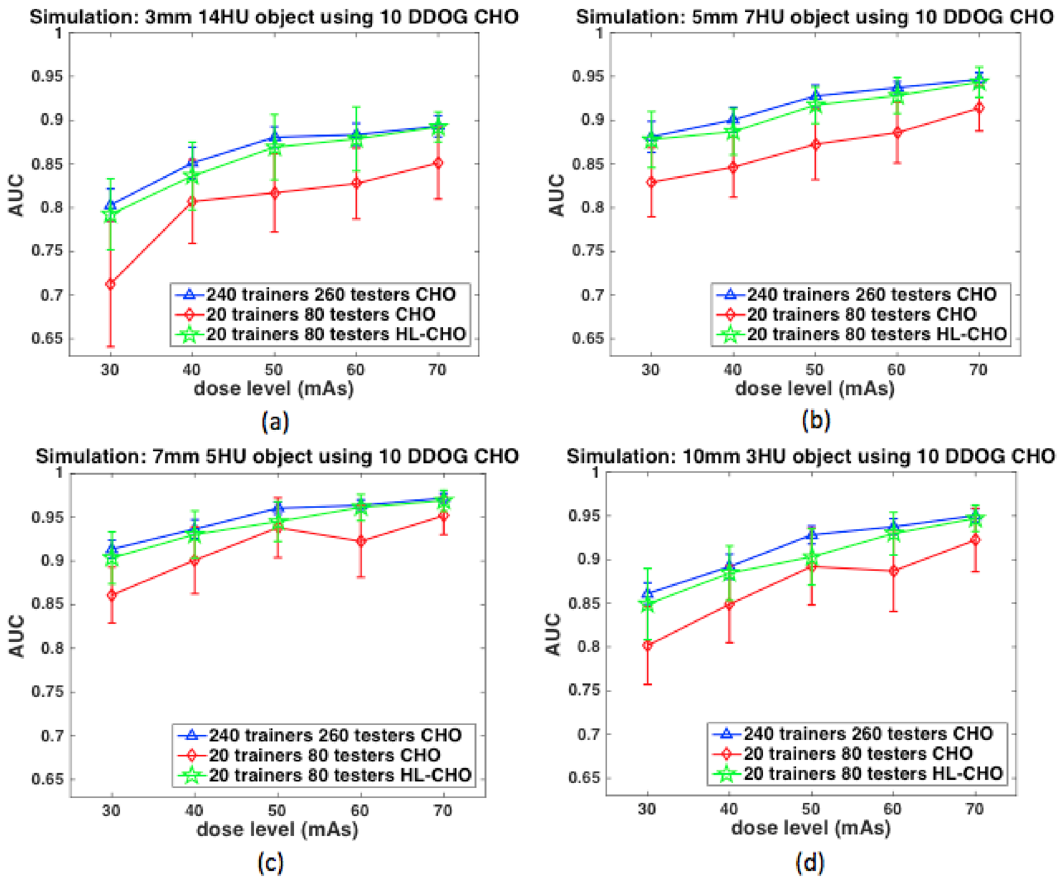

Figure 6.4: Simulation study. The comparison of AUC values for 500 image pairs CHO, 100 image pairs CHO, and 100 image pairs HL-CHO using simulated images.(a) The object of 3 mm 14 HU (b) The object of 5 mm 7 HU (c) The object of 7 mm 5 HU (d) The object of 10 mm 3 HU.

were considered: (i) the baseline, traditional method with Full data (ii) traditional method with less data (iii) proposed method (HL-CHO/HL-CSLO) with less data. Full data refers using 240 samples for training and 260 for testing. Less data means that there are only 20 samples for training and 80 samples for testing. All the images were real CT images. The FOM for the pure detection task and the combination of detection and estimation task is AUC and EAUC value respectively. The mean and variance of FOM at each dose level is estimated by the shuffle method (Fukunaga and Hayes (1989)) using 20 different shuffles of the data. The accuracy is defined by the ratio of mean value of the less data to the mean value of the baseline. The X-Ray current range was chosen such that the FOM value was neither close to 1 nor close to the guessing value. FOM value equals to 1 means that the observer can easily detect if the signal/tumor exists and distinguish the physical properties of signal/tumor. The guessing value will be described later.

For the pure detection task, the MITA CCT 183 phantom was used. With ROC analysis the guessing observer has an AUC of 0.5. The X-Ray current range is between 90 mAs and 140 mAs, and the accuracy is 85% to 92% and 97% to 100% for CHO and HL-CHO with less data respectively. The plots of comparisons are shown in figure 6.5. The results of the pure detection task indicate that our proposed approach can efficiently reduce the required data without losing accuracy.

For the detection and size estimation task, X-Ray current operation range was selected between 60 mAs to 120 mAs. The blind guessing value is $\frac{1}{2} \times \frac{1}{\text{the number of estimated physical properties}}$. Again the $\frac{1}{2}$ comes from detection by guessing. The number of estimate physical properties, in this case, is 4. So the blind guessing value is 0.125. The QRM-LC-FD1 phantom was used for this task. In this task, the observer not only had to distinguish if the signal exists but also estimate the size of the signal. The results were shown in figure 6.6. In this task, the accuracy is 84%$\sim$88% and 97%$\sim$99% for CSLO and HL-CSLO method respectively. Based on the results of this study, the proposed approach only needs 20% of the full dataset to reach about the same performance of the traditional approach with the full dataset.

For the detection and contrast estimation task, the phantom QRM-LC-FD4 was used for this task. Similarly to the task of combined detection and size estimation, there were also 4 objects of same size with 4 different contrasts so the blind guessing value is also 0.125. The operation range of X-Ray current was 15 mAs 80 mAs. The observer, in this task, had to tell if there is a signal in the testing image or not and determined the contrast of the signal. Results show that the accuracy is 92%∼96% for CSLO and 97%∼100% for the HL-CSLO. The plot of EAUC vs. radiation dose level is shown in figure 6.7. For the CSLO approach, the accuracy in the combined detection and contrast estimation task is better than the accuracy in the combined detection and size estimation task due to the contrasts of objects in QRM-LC-FD4 phantom were higher than the contrasts of objects in QRM-LC-FD1 phantom. This causes the accuracy to become better in the lower dose range. As with the previous task, the proposed approach still maintains higher performance than the traditional approach under this task.

In the study of combination of detection, size and contrast estimation, the observer had to not only decide whether the signal is in the image or not but also make decisions on the size and contrast of the detected signal. The blind guessing value in this complicated case is 0.0625 according to the guessing value equation discussed above and the number of estimate physical properties.The operated X-Ray current was from 80 mAs to 160 mAs for MITA CCT 189 phantom. The accuracy of CSLO method is 88% to 91% and the accuracy of HL-CSLO is 94% to 99%. The results are shown in figure 6.8. For the CLSO approach with less data, the accuracy in this task is slightly higher than the combination task of the detection and size estimation. This does not imply that CSLO approach with less data in the combination task of detection, size and contrast estimation has better performance because the background of phantom (MITA CCT 189) used in this task was different. The proposed approach, HL-CSLO, again keeps high accuracy even in the low radiation dose level range.

The accuracies for all the tasks performed by traditional method and proposed method are listed on the Table 6.1. All the studies were based on 80% data reduction. Overall, the traditional approaches, CHO and CSLO, both suffer while

Table 6.1: Accuracy Summary of 80% data reduction

	Traditional Method	Proposed Method
Detection Task	85.88%-91.16%	97.38%-99.86%
Detection and Size Estimation Task	84.43%-87.51%	97.13%-98.85%
Detection and Contrast Estimation Task	92.84%-95.28%	97.46%-99.69%
Detection, Size and Contrast Estimation Task	88.17%-90.65%	94.80%-98.33%

the data reduced especially in the case of lower radiation dose level. The proposed approaches, HL-CHO and HL-CSLO, have stable performance with high accuracy for all for all radiation dose level s in all tasks.

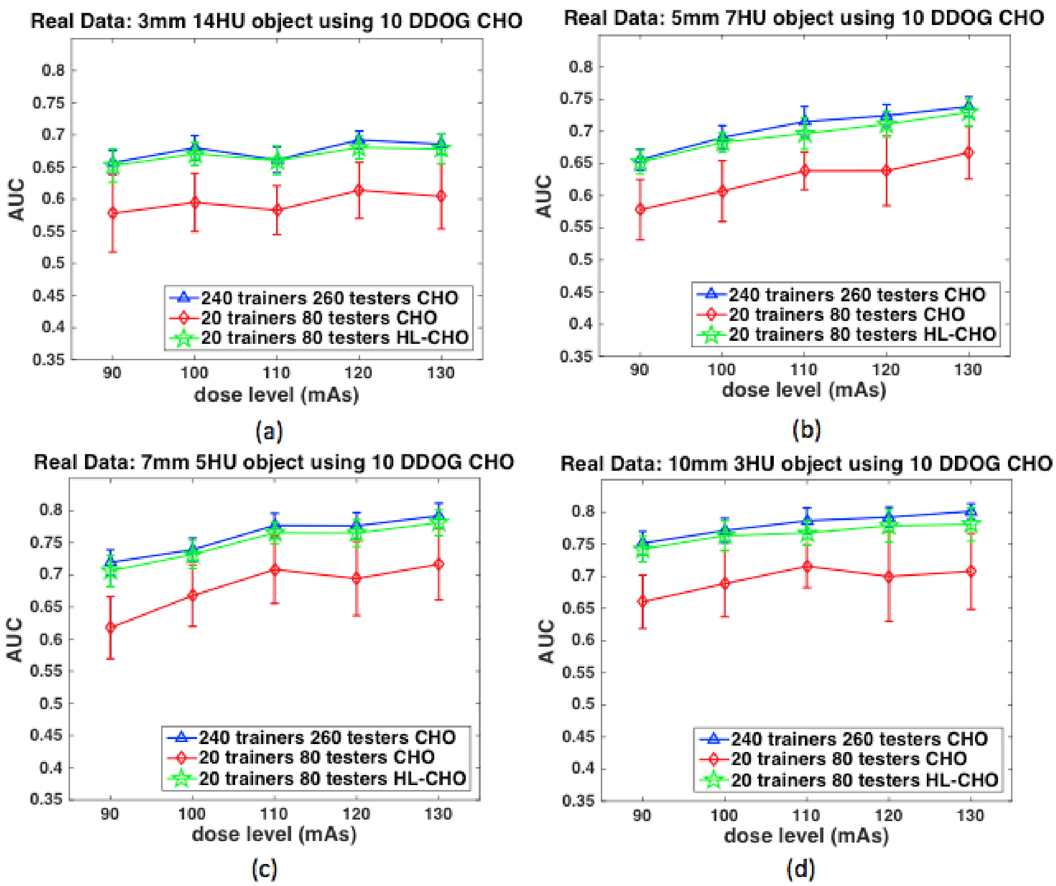

Figure 6.5: The detection task. The comparison of AUC values for 500 image pairs CHO, 100 image pairs CHO, and 100 image pairs HL-CHO from real image data sets. (a) The object of 3 mm 14 HU (b) The object of 5 mm 7 HU (c) The object of 7 mm 5 HU (d) The object of 10 mm 3 HU.

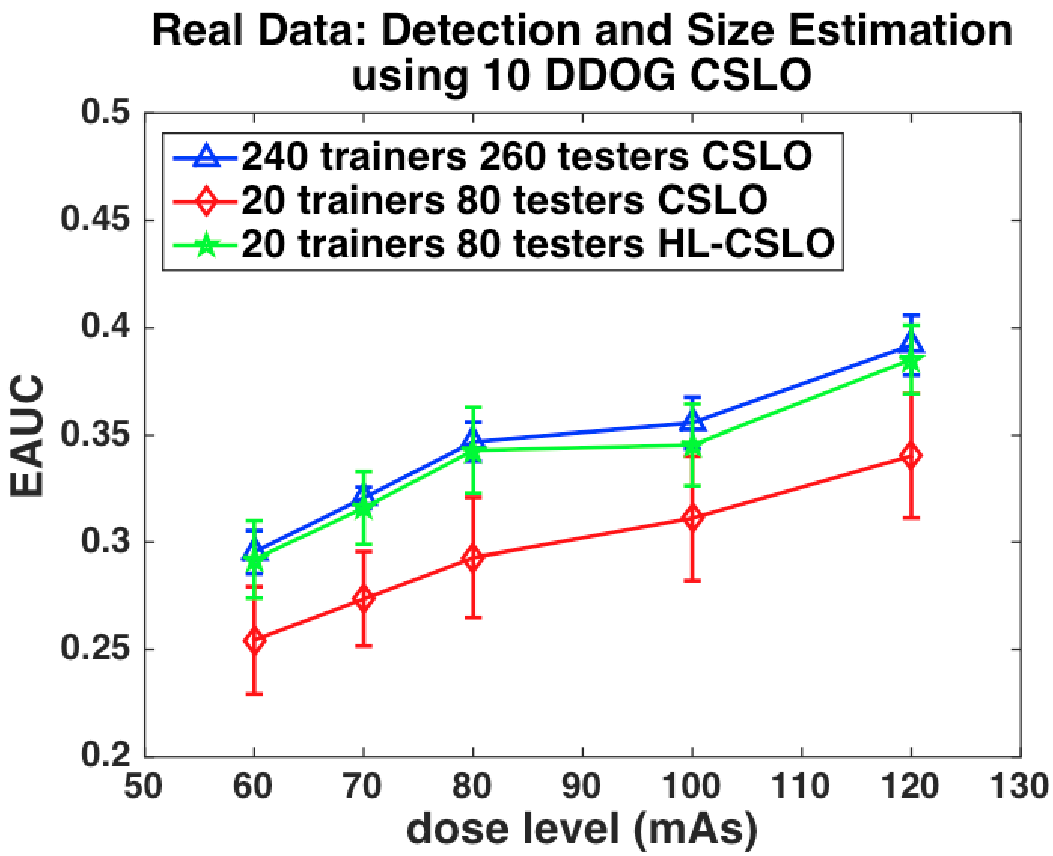

Figure 6.6: The detection and size estimation task. Comparisions of EAUC values for 500 image pairs CSLO, 100 image pairs CSLO, and 100 image pairs HL-CSLO from real image data sets.

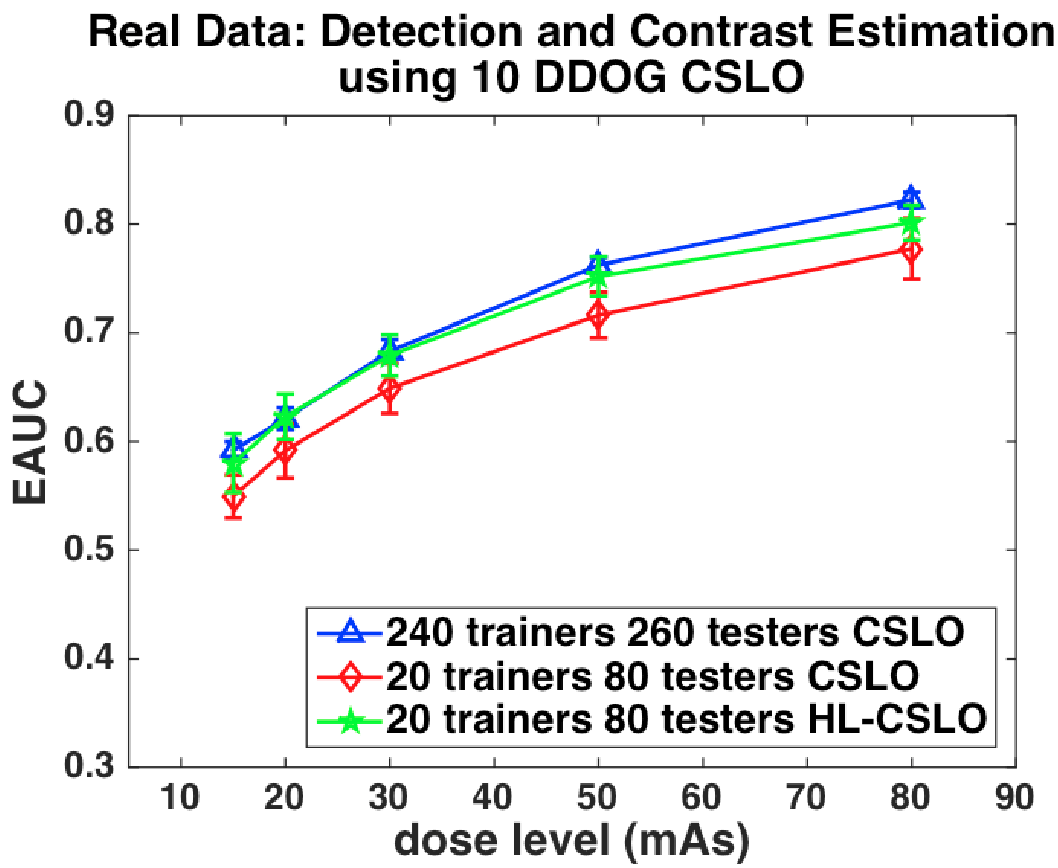

Figure 6.7: The detection and contrast estimation task. Comparisions of EAUC values for 500 image pairs CSLO, 100 image pairs CSLO, and 100 image pairs HL-CSLO from real image data sets.

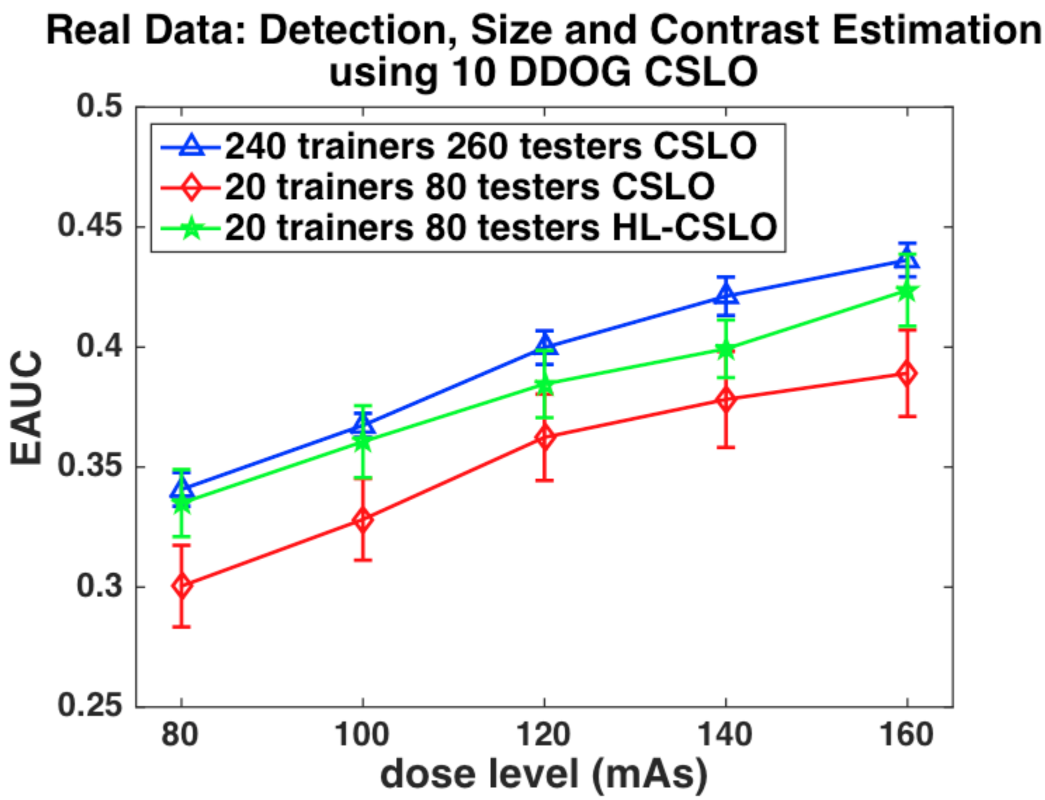

Figure 6.8: The combined detection and size as well as contrast estimation task. Comparisions of EAUC values for 500 image pairs CSLO, 100 image pairs CSLO, and 100 image pairs HL-CSLO from real image data sets.

CHAPTER 7

Conclusions

7.1 Summary

As mentioned in the first chapter, although different image quality metrics such as modulation transfer function (MTF) and noise power spectrum (NPS) have been used in CT, tasked-based model observers were selected in this work to study the image quality performance of reconstruction algorithms for CT imaging. In the concepts of MTF and NPS, the assumptions of linearity in the reconstruction algorithms and shift-invariance of the noise are not valid for IR in CT system because linearity is not used in IR and the divergent X-Ray violates the property of shift invariance. The model observers require an accurate covariance matrix in order to calculate observer performance. To have a good estimate of the covariance matrix, it is important to get reliable statistical information from image samples by using repeated CT scans. We did this by scanning the phantoms using exactly the same settings multiple times for both head and body modes. It should be noted that axial correlations might exist. Ideally, the experiment should be designed to scan the phantom 500 times to get 500 images for one dose level. However, this idea is not practical. Take the study of a pure detection task as an example, there were 21 dose levels in head mode and 34 dose levels in body mode. Ideally, this would require 27500 total CT acquisitions. For our study, we obtained 2750 acquisitions and avoided axial correlations by randomly selecting the noise ROIs from different locations.

In chapter 2, the X-Ray radiation dose level required for low contrast object detection in head and body phantoms was examined and derived quantitatively. The work presented here provides an objective and robust way to quantitatively assess the image quality impact of newly introduced CT dose saving features for the pure detection task. For the selection of channels, both DDOG and Gabor channels are well known channels that mimic human visual system. In our study, we showed that these two channels achieved similar results. However, we selected DDOG channels because the number of Gabor channels, in general, is larger. To have good estimate of the channelized covariance matrix, more training images may be required. According to our results, the performance of the new IR algorithm shows that the image quality has been greatly improved compared to the traditional FBP algorithm for both location known detection and location unknown detection tasks. Thus, to achieve a FBP-equivalent image quality, the iterative reconstruction algorithm developed by GE can operate with up to 82% dose reduction.

In chapter 3, we used the CSLO to measure the image quality under a fuller picture of the task, detection and estimation, at different dose levels for two different reconstruction methods. Radiation dose levels required for low contrast object detection and estimation in body phantoms were examined and evaluated quantitatively. The work presented here provides an objective way to assess the image quality of CT imaging systems for dose saving evaluation. Unlike pure detection tasks, there is currently no proper way of adding internal noise in the combination of detection and estimation tasks so internal noise was not used in this study. Our results provide upper bounds for observer studies. The dose reduction was 50% for both signal- location-known detection/size estimation and signal- location-known detection/contrast estimation task. Note here, although the capability of dose reduction (54%) from IR under the hybrid task of signal location known detection, size estimation and contrast estimation is higher than the dose reduction from combined signal-location-known detection/size estimation and signal-location-known detection/contrast estimation, it is hard to tell if the IR has more superiority in the more difficult task because the phantom used in this task was different than the other two tasks. A similar explanation could also be applied on the signal-location-unknown study.

A number of approaches have been used in the effort to minimize radiation dose used in CT. These include X-Ray tube current modulation, optimized tube voltage, etc. In chapter 2 and 3, we have evaluated the performance of the IR algorithm developed by GE and our results indicate that the IR algorithm can be another tool allowing for a reduction of radiation dose. According to our results, the performance of the new IR algorithm shows that the image quality has been greatly improved compared to the traditional FBP algorithm under different tasks. This indicates that we can achieve a FBP similar performance using the IR algorithm but at a lower exposure and, hence, lower dose to the patient.

In chapter 4, We studied the performances of different protocols under contrast enhancement. Iodine attenuation increases as the tube voltage decreases. The physics behind this phenomenon is that the energies in the X-Ray shifted closer to the K-absorption edge of iodine. Thus, decreasing the tube potential leads to an increase in the iodine contrast. However, this beneficial effect is compensated by the increased attenuation in tissue that causes higher noise. Hence, to optimize image quality, it is important to know which protocol provides better performance.

When the $CTDI_{vol}$ is kept the same for all protocols, it is obvious that the protocol using lower voltage gives better image quality based on our understanding of CNR. Although a physics-based image quality metric such as CNR could intuitively give us a general idea of the image quality of a system, it is limited as an assessment of the image quality in a clinical sense. To establish a bridge between radiologists and medical physicists, and therefore between the clinical and physical image quality, task-based measurements are currently suggested. In our studies, tasks related to clinical studies such as detection tasks, estimation tasks, and combination tasks of detection and estimation were all evaluated by mathematical model observers. Results show that the image quality not only depends upon contrasts of objects but also depends upon sizes of objects, and the sizes of phantoms.

Although our simulation images were all generated based on real images, our results here can only provide some ideas of how model observers can help to optimize the protocols because of the limited data. Here, the signals were simulated based on the real signals at one dose level, however, the shapes of signals might be changed for different protocols at different dose levels. Besides, the shape of the NPS and the frequency associated with the peak of NPS are also varied as the radiation dose is decreased. To understand the performances of different protocols at specific dose ranges, real images generated from that specific dose ranges are required.

In chapter 5, a task-based image quality evaluation using a threshold-based observer has been designed and performed to evaluate the temporal resolution in Cardiac CT imaging systems. A motion phantom filled with water accompanied with a 3 mm vessel phantom was used for simulating the cardiac motion. Cases of different heart-beating rates, reconstruction phases, and image reconstruction algorithms were compared. In the case of 60 bpm, improvements could be observed on the estimated value of 52.68% when the phase was selected at the optimal phase. However, the image quality improvements are barely seen when SSF was applied on the images. From our analysis, the image quality of the 80 bpm case was improved greatly when either the phase was selected at optimal phase or the images were generated by SSF. The image quality of 80 bpm case, when using SSF, reached a similar level as the 60 bpm case. This implies that the image reconstruction algorithm, SSF, plays a more important role in image quality improvement in this study. Overall, superior reconstruction phase and SSF both can improve the image quality for higher heart-beating rates. The improvement offered by SSF is much more significant than that of phase selection, especially in the case of higher heart-beating rates. A more complicated background including the internal structural noise will be considered in a future study. Quantitatively speaking, our results suggest that SSF is a more promising reconstruction algorithm than the traditional FBP algorithm in Cardiac CT systems, particularly when a faster beating heart is scanned.

In chapter 6 where both the simulated and the real data were used, we found that the proposed methods can match the ground truth with high accuracy while using much less data. Instead of using 500 image (240 training and 260 testing) pairs, the new approach only requires 100 image pairs (20 training and 80 testing) for every dose level for all tasks. This implies that instead of 50 scans for every radiation dose level, we only need 10 scans-a substantial reduction. The 40 images generated by high dose levels can be repeatedly used for different dose levels. The amount of image data required can be reduced up to 80% with our proposed method. Thus, the usage of model observers to assess image quality can become a practical tool for routine evaluations of CT imaging systems.

The images from high dose efficiently estimate the information of the signal such as modulation transfer function (MTF) with much less data because of the higher signal to noise ratio in the image data. Note this property is only valid for images from linear reconstruction algorithm such as FBP. For iterative reconstruction images, the signal estimated by high dose will be an approximation because the noise background is nonlinear. The channelized covariance matrix estimated by the process of maximizing the Leave-One-Out Likelihood (LOOL) is more accurate because the covariance matrix estimated by the LOOL mixes the sample channelized covariance matrices and common channelized covariance matrices. This hybrid of matrices makes the inverse of matrix more stable with a sample size closer to the dimension of the channelized vector. The incorporation of channelized estimated signals from high dose image and channelized-LOOL covariance matrices makes the FOM value very close to the baseline.

Our results in chapter 6 consistently indicate that the data can be reduced up to 80% if our proposed method is applied on the CHO and CSLO. Thus, for both the pure detection and the combination of detection and estimation tasks, the proposed methods, HL-CHO and HL-CSLO, can be used to quantify the dose saving capability as well as the low contrast detectability (LCD) performance of a CT system. Under controlled imaging conditions, HL-CHO and HL-CSLO can be used to compare the dose and LCD performance of different algorithms and components on one system. They can also be used as a tool for the acceptance test for a newly developed iterative reconstruction method. In summary, HL-CHO and HL-CSLO are good candidates for the regular quality assurance and quality control tests in hospitals as well.

7.2 Future Work

In this work, internal noise was not considered in any task. Although approaches to adding internal noise in the pure detection task have been addressed in the literature, the current methods for adding internal noise for the estimation task is not clear. Our results here provide an upper bound of the observer performance. To have more realistic results, further work such as adding internal noise is needed. However, there are two problems of current approaches of adding internal noise. First, they all require human observer studies for internal noise constant calibrations, which are very time consuming. Secondly, to my best understanding of these approaches, the internal noise constant is calibrated based on only one size of the object and one dose level. The first part of my future work will be deriving a function for internal noise with parameters of the object size and the radiation dose level and designing an approach that could give the users of model observers a reasonable numerical range of internal noise constants instead of requiring calibrations from human observers.

One of the limitations of the current phantom design is that the background is uniform. In reality, the background is very complex. To make the image quality measurements closer to realistic situations, adding noise with the same frequency of background structures into channels will be considered. Another way to make backgrounds more realistic is to use anatomic images (real patient images) with lesions inserted into projection data before image reconstruction process. Current results for the efficiencies from model observers are higher than those of human observers, adding noise or using anatomic backgrounds will definitely degrade the performance of model observers to match the performance of human observers (Barrett et al. (2015b); Young et al. (2013)). Another limitation is that the current signals/tumors embedded in the phantoms are all circularly symmetric in shape. Most of the channels for channelized model observers are designed for circular signals. Designing a channelized model observer that could efficiently extract the information of signals of irregular shape is also desired in the second part of the future plan.

The third part of the future work is related to dose reduction. In chapter 2 and 3, we have already seen that IR algorithms efficiently make huge differences in image quality improvements compared to the traditional FBP for different tasks. However, current techniques of IR algorithms do not combine model observers inside the iteration loops. Practical integration of the model observers into IR in CT could make IR function more efficiently. Hence, designing a IR with a closed-loop model observers or model observer based IR (MOIR) is challenging and will be considered in the future study as well. In addition to using software such as iterative algorithms to reduce radiation dose to the patients, hardware such as auto exposure control (AEC) is also a popular way to achieve the dose reduction. The key point will be how to optimally use the data from the scout scan to generate the profile for the X-Ray tube current. For implementation, applying model observers on the sinogram data could help us in the optimization. The benefit of this method is that it requires less data than methods of using model observers scanned on image domain data.

Finally, in addition to CT, there are other modalities used in clinical studies. For different modalities, the image backgrounds are different. Besides, the amount of image data might not be sufficient to get good statistics in PET/CT and SPECT/CT. To select or design model observers that can correctly and efficiently evaluate image quality is very important. It will be of interest to see if current approaches presented in this dissertation can be applied in different modalities such as PET/CT and SPECT/CT system.

REFERENCES

Abbey, C. K. and H. H. Barrett (2001). Human-and-model-observer performance in ramp-spectrum noise: Effects of regularization and object variability. *J Opt Soc Am A*, **18**.

Abbey, C. K. and J. M. Boone (2008). An ideal observer for a model of x-ray imaging in breast parenchymal tissue. *Lecture Notes in Computer Science*, **5116**, pp. 393–400.

Abbey, C. K., M. P. Eckstein, S. S. Shimozaki, A. H. Baydush, D. M. Catarious, and C. E. Floyd (2002). Human-observer templates for detection of a simulated lesion in mammographic images. *Proc. SPIE*, **4686**.

Achenbach, S., D. Ropers, J. Holle, G. Muschiol, W. G. Daniel, and W. Moshage (2000). In-plane coronary arterial motion velocity: measurement with electron beam CT. *Radiology*, **216**.

Barrett, H. H., C. K. Abbey, and E. Clarkson (1998). Objective assessment of image quality. III. ROC metric, ideal observers, and likelihood-generating functions. *J. Opt. Soc. Am. A*, **15**.

Barrett, H. H., L. R. Furenlid, J. Y. Hesterman, M. A. Kupinski, and M. K. Whitaker (2008). Adaptive SPECT. *IEEE Trans. Med. Imaging*, **27**.

Barrett, H. H., M. A. Kupinski, and E. Clarkson (2005). Probabilistic foundations of the MRMC method,? in: MP Eckstein, Y Jiang, eds. Medical imaging 2005: Image perception, observer performance, and technology assessment. *Proc. SPIE*, **5749**.

Barrett, H. H. and K. J. Myers (2003). Foundations of image science. *John Wiley and Sons*.

Barrett, H. H., K. J. Myers, C. Hoeschen, M. A. Kupinski, and M. P. Little (2015a). Task-based measures of image quality and their relation to radiation dose and patient risk. *Phys. Med. Biol.*, **60**.

Barrett, H. H., K. J. Myers, C. Hoeschen, M. A. Kupinski, and M. P. Little (2015b). Task-based measures of image quality and their relation to radiation dose and patient risk. *Phys. Med. Biol.*, **60**.

Bhaglia, R., J. Pack, J. Miller, and M. Iatrou (2012). Nonrigid registration-based coronary artery motion correction for cardiac computed tomography. *Med. Phys.*, **39**.

Biden, S. V., R. F. Wagner, and G. Campbell (2000). Components-of-variance models and multiple-bootstrap experiments: An alternative method for random-effects, receiver operating characteristic analysis. *Acad. Radiol.*, **7**.

Boone, J. M. (2001). Determination of the presampled MTF in computed tomography. *Med Phys*, **28**, pp. 356–360.

Boyd, D. and M. Lipton (1983). Cardiac computed tomography. *Proc. IEEE*, **71**.

Burgess, A. E. (2011). Visual perception studies and observer models in medical imaging. *Semi Nucl Med*, **41**, pp. 419–436.

Chen, G. H., J. Tang, and J. Hsieh (2009). Temporal resolution improvement using prior image constrained compressed sensing PICCS. *Med. Phys.*, **36**.

Chen, G. H., J. Tang, and J. Hsieh (2010). Temporal resolution improvement using prior image constrained compressed sensing PICCS. *Med. Phys.*, **36**.

Cho, J., S. Ramani, and J. Fessler (2012). Motion-compensated image reconstruction with alternating minimization. *The second international conference on image formation in X-ray computed tomography.*

Chun, S. and J. Fessler (2009). Joint image reconstruction and nonrigid motion estimation with a simple penalty that encourages local invertibility. *Proc. of SPIE. Medical Imaging: Image Processing*, **7258**.

Clarkson, E. (2007). Estimation receiver operating characteristic curve and ideal observers for combined detection/estimation tasks. *J Opt Soc Am A Opt Image Sci Vis*, **24**.

Das, M., H. C. Gifford, M. O'Connor, and S. J. Glick (2011). Penalized maximum likelihood reconstruction for improved microcalcification detection in breast tomosynthesis. *IEEE Trans. Med. Imaging*, **30**.

Detrano, R. C., M. Anderson, J. Nelson, N. Wong, J. J. Carr, M. McNittGray, and D. E. Bild (2005). Coronary calcium measurements: Effect of CT scanner type and calcium measure on rescan reproducibility-MESA study. *Radiology*, **236**, pp. 477–484.

Dobbin, J. T., E. Samei, N. T. Ranger, and Y. Chen (2006). Intercomparison of methods for image quality characterization. II. Noise power spectrum. *Med. Phys.*, **33**.

Dorfman, D. D., K. S. Berbaum, R. V. Lenth, Y. F. Chen, and B. A. Donaghy (1998). Monte Carlo validation of a multireader method for receiver operating characteristic discrete rating data: Factorial experimental design. *Acad. Radiol.*, **5**.

Dorfman, D. D., K. S. Berbaum, and C. E. Metz (1992). Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method. *Invest. Radiol.*, **27**.

Efron, B. and R. J. Tibshirani (1993). An Introduction to the Bootstrap: Monographs on Statistics and Applied Probability. *Chapman and Hall New York NY.*

Feigal, D. W. J. (2002). FDA public health notification: reducing radiation risk from computed tomography for pediatric and small adult patients. *Int J Trauma Nurs*, **8**, pp. 1–2.

Flohr, T., C. McCollough, H. Bruder, M. Petersilka, K. Gruber, C. Suss, M. Grasruck, K. Stierstorfer, B. Krauss, R. R., A. N. Primak, A. Kuttner, S. Achenbach, C. Becker, A. Kopp, and B. M. Ohnesorge (2006). First performance evaluation of a dual-source CT (DSCT) system. *Eur. Radiology*, **16**.

Fukunaga, K. and R. R. Hayes (1989). Effects of sample size in classifier design. *IEEE Trans. on Pattern Analysis and Machine Learning*, **11**, pp. 873–885.

Gallas, B. D. (2006). One-Shot Estimate of MRMC Variance: AUC. *Acad. Radiol.*, **13**.

Gallas, B. D. and H. H. Barrett (2003). Validating the use of channels to estimate the ideal linear observer. *J Opt Soc Am A*, **20**.

Geisler, W. S. (2003). Ideal observer analysis. The visual neurosciences. *Boston: MIT press*, pp. 771–781.

Geisler, W. S. (2011). Contributions of ideal observer theory to vision research. *Vision Res*, **51**, pp. 771–781.

Gifford, H. C., M. A. King, D. J. de Vries, and E. J. Soares (2000a). Channelized Hotelling and human observer correlation for lesion detection in hepatic SPECT imaging. *J. Nucl. Med.*, **41**.

Gifford, H. C., M. A. King, R. G. Wells, W. G. Hawkins, M. V. Narayanan, and P. H. Pretorius (2000b). LROC analysis of detector-response compensation in SPECT. *IEEE Trans. Med. Imaging*, **19**.

Graff, C. G. and K. J. Myers (2011). The ideal observer objective assessment metric for magnetic resonance imaging: application to signal detection task. *Inf. Process Med. Imaging*, **22**, pp. 760–771.

Green, D. M. and J. A. Swets (1966). Signal detection theory and psychophysics. *John Wiley and Sons*.

Gross, K., M. A. Kupinski, T. E. Peterson, and E. Clarkson (2003). Optimization a multiple-pinhole SPECT system using the ideal observer. *Proc. SPIE*, **5034**.

Hara, A., R. G. Paden, A. C. Silva, J. Kujak, and H. J. Lawder (2009). Iterative reconstruction technique for reducing body radiation dose at CT: feasibility study. *Am J. Roentgenol*, **193**, pp. 764–771.

He, X., B. S. Caffo, and E. C. Frey (2008). Toward realistic and practical ideal observer (IO) estimation for the optimization of medical imaging systems. *IEEE Trans. on Med. Imaging*, **27**, pp. 1535–1543.

Herman, G. T. (1980). The Fundamentals of Computed Tomography. *Academic Press New York*.

Hoffbeck, J. P. and D. A. Landgrebe (1996). Covariance Matrix Estimation and Classification with Limited Training Data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **18**.

Horiguchi, J., T. Nakanishi, A. Tamura, and K. Ito (2002). Coronary artery calcium scoring using multicardiac computed tomography. *J. Comput. Assist. Tomogr.*, **26**, pp. 880–885.

Hsieh, J., B. Nett, Z. Yu, K. Sauer, J. Thibault, and C. Bouman (2013). Recent Advances in CT Image Reconstruction. *Current Radiology Reports*.

Hsieh, J. and R. Senzig (2002). Dual cardiac CT scanner. *United States Patent*.

Iatrou, M., J. Pack, R. Bhagalia, D. Beque, and J. Seamans (2010). Coronary artery motion estimation and compensation: A feasibility study. *Nuclear Science Symposium Conference Record (NSS/MIC) IEEE*.

International Commission on Radiation Units and Meausrements. ICRU Report No.87 (2012). Radiation dose and image-quality assessment in computed tomography. *J ICRU*, **12**, pp. 1–149.

Jakobs, T. F. and C. R. Becker (2002). Multislice helical CT of the heart with retrospective ECG gating: Reduction of radiation exposure by ECG-controlled tube current modulation. *Eur. Radiol*, **12**, pp. 1081–1086.

Johnson, T. R., K. Nikolaou, B. J. Wintersperger, A. W. Leber, F. Ziegler von, M. Rist, C. R. Becker, G. Steinbeck, and P. Boekstegers (2006). Dual-source CT cardiac imaging: Initial experience. *Eur. Radiol.*, **16**, pp. 1409–1415.

Judy, P. F. (1976). The line spread function and modulation transfer function of a computed tomographic scanner. *Med Phys*, **3**, pp. 233–236.

KachelrieB, M., M. Knaup, and W. A. Kalender (2006). Multi-threaded cardiac CT. *Med. Phys.*, **33**.

Kadrmas, D. J. and P. E. Christian (2002). Comparative evaluation of lesion detectability for 6 PET imaging using a highly reproducible whole-body phantom with Na lesions and location ROC analysis. *J. Nucl. Med.*, **43**.

Kak, A. C. and M. Slaney (1988). Principles of Computerized Tomographic Imaging. *IEEE*.

Kalayeh, H. M. and D. A. Landgrebe (1983). Predicting the Required Number of Training Samples. *Current Radiology Reports*.

Kalender, W. A., H. Wolf, and C. Suess (1999). Dose reduction in CT by anatomically adapted tube current modulation: Phantom measurements. *Med. Phys.*, **26**, pp. 2248–2253.

Kallender, W. A., W. Seissler, and P. Vock (1989). Single-breath-hold spiral volumetric CT by continuous patient translation and scanner rotation. *Radiology*, **173**.

Kalra, M. K., M. M. Maher, D. V. Sahani, M. A. Blake, P. F. Hahn, G. B. Avinash, T. L. Toth, E. Halperm, and S. Saini (2003). Low-dose CT of the abdomen: Evaluation of image improvement with use of noise reduction filters-pilot study. *Radiology*, **228**, pp. 251–256.

Khurd, P. and G. Gindi (2005). Decision strategies that maximize the area under the LROC curve. *IEEE Trans. Med. Imaging*, **24**.

Krupinski, E. A. (1996). Visual scanning patterns of radiologists searching mammograms. *Acad. Radiol.*, **3**.

Kupinski, M. A., J. W. Hopppin, E. Clarkson, and H. H. Barrett (2003). Ideal-observer computation in medical imaging with use of Markov-chain Monte Carlo techniques. *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.*, **20**, pp. 430–438.

Leber, A. W., A. Knez, C. Becker, C. Becker, F. Ziegler von, K. Nikolaou, C. Rist, M. Reiser, C. White, G. Steinbeck, and P. Boekstegers (2004). Accuracy of multidetector spiral computed tomography in identifying and differentiating the composition of coronary atherosclerotic plaques. *J. Am. Coll. Cardiol*, **43**, pp. 1241–1247.

Leber, A. W., A. Knez, F. Ziegler von, A. Becker, K. Nikolaou, S. Paul, B. Wintersperger, M. Reiser, C. R. Becker, G. Steinbeck, and P. Boekstegers (2005). Quantification of obstructive and nonobstructive coronary lesions by 64-slice computed tomography: A comparative study with quantitative coronary angiography and intravascular ultrasound. *J. Am. Coll. Cardiol.*, **46**, pp. 147–154.

Lee, C. J., M. A. Kupinski, and L. Volokh (2013). Assessment of cardiac single-photon emission computed tomography performance using a scanning linear observer. *Med. Phys.*, **40**, p. 011906.

Mao, S., M. J. Budoff, H. Bakhsheshi, and S. C. Liu (2001). Improved reproducibility of coronary artery calcium scoring by electron beam tomography with a new electrocar-diographic trigger method. *Invest. Radiol.*, **36**, pp. 363–367.

Marin, D., R. C. Nelson, S. T. Schindera, S. Richard, and R. S. Youngblood (2010). Low-tube-voltage, high-tube-current multidetector abdominal CT: Improved image quality and decreased radiation dose with adaptive statistical iterative reconstruction algorithm-initial clinical experience. *Radiology*, **254**, pp. 145–153.

McCollough, C. H., M. R. Bruesewitz, and J. M. J. Kofler (2006). CT dose reduction and dose management tools: Overview of available options. *Radiographics*, **26**, pp. 503–512.

Mettler, F. A. J., M. Bhargavan, and K. Faulkner (2009). Radiologic and nuclear medicine studies in the United States and worldwide: Frequency, radiation dose, and comparison with other radiation sources-1950-2007. *Radiology*, **253**, pp. 520–531.

Myers, K. J. and H. H. Barrett (1987). Addition of A Channel Mechanism to the Ideal-Observer Model. *J Opt Soc Am A Opt Image Sci Vis*, **4**.

Nakayama, Y., K. Awai, Y. Funama, M. Hatemura, M. Imuta, T. Nakaura, D. Ryu, S. Morishita, S. Sultana, N. Sato, and Y. Yamashita (2005). Abdominal CT with low tube voltage: preliminary observations about radiation dose, contrast enhancement, image quality, and noise. *Radiology*, **237**, pp. 945–951.

Nikoloff, E. L. and R. Wagner (1985). SNR and DQE analysis of broad spectrum X-ray imaging. *Phys Med Biol*, **30**, p. 519.

Nishimura, H. and O. Miyazaki (1988). CT system for spirally scanning subject on a movable bed synchronized to x-ray tube revolution. *U.S. Patent*.

Novak, L. (1984). On the sensitivity of Bayes and Fisher classifier in radar target detection. *Proc. 18th Asilomar Conf. Circuits*, pp. 5–7.

Obuchowski, N. A. (1995). Multireader receiver operating characteristic curve studies: Hypothesis testing and sample size estimation using an analysis of variance approach with dependent observations. *Acad. Radiol.*, **2**.

Obuchowski, N. A. and H. E. Rockette (1995). Hypothesis testing of diagnostic accuracy for multiple readers and multiple tests: An ANOVA approach with dependent observations. *Commun. Sat.*, **24**.

Oliver, M. F., E. Samuel, P. Morley, G. P. Young, and P. L. Kapur (1964). Detection of coronary artery calcification during life. *Lancet*, **1**.

Pack, J. and B. Claus (2011). An analysis of motion artifacts in CT and implications for motion compensation. *Proc. of the 2nd International Conf. on Image Formation in X-Ray Computed Tomography.*

Petersilka, M., H. Bruder, B. Krauss, K. Stierstorfer, and T. G. Flohr (2008). Technical principles of dual source CT. *European Journal of Radiology*, **68**.

Raff, G. L., M. J. Gallagher, W. W. O'Neill, and J. A. Goldstein (2005). Diagnostic accuracy of noninvasive coronary angiography using 64-slice spiral computed tomography. *J. Am. Coll. Cardiol*, **46**, pp. 552–557.

Roberts, H. C., T. P. Roberts, T. Y. Lee, and W. P. Dillow (2002). Dynamic, Contrast-Enhanced CT of Human Brain Tumors: Quantitative Assessment of Blood Volume, Blood Flow, and Microvascular Permeability: Report of Two Cases. *AJNR Am J. Neuroradiol*, **23**.

Roe, C. A. and C. E. Metz (1997a). Dorfamn-Berbaum-Metz method for statistical analysis of multireader, multimodality receiver operating characteristic (ROC) data: Validation with computer simulation. *Acad. Radiol.*, **4**.

Roe, C. A. and C. E. Metz (1997b). Variance-component modeling in the analysis of receiver operating characteristic (ROC) index estimates. *Acad. Radiol.*, **4**.

Rohkohl, C., H. Bruder, K. Stierstorfer, and T. Flohr (2012). Improving best-phase image quality in cardiac CT by motion correction with MAM optimization. *The second international conference on image formation in X-ray computed tomography.*

Rolland, J. P. and H. H. Barrett (1992). Effect of random background inhomogeneity on observer detection performance. *J Opt Soc Am A Opt Image Sci Vis*, **9**.

Ropers, D., U. Baum, K. Pohle, K. Anders, S. Ulzheimer, B. Ohnesorge, C. Schlundt, W. Bautz, W. Daniel, and S. Achenbach (2003). Detection of coronary artery stenoses with thin-slice multi-detector row spiral computed tomography and multiplanar reconstruction. *Circulation*, **107**, pp. 664–666.

Schondube, H., T. Allmendinger, K. Stierstorfer, H. Bruder, and T. Flohr (2012). Evaluation of a novel CT image reconstruction algorithm with enhanced temporal resolution. *Proc. SPIE*, **7916**.

Schroeder, S., S. Achenbach, F. Bengel, C. Burgstahler, F. Cademartir, P. de Feyter, R. George, P. Kaufmann, A. Kopp, J. Knuuti, D. Ropers, J. Schuijf, L. Tops, and J. Bax (2007). Cardiac computed tomography: indications, applications, limitations and training requirements. *European Heart Journal*.

Seibert, J. A. (2004). X-Ray imaging physics for nuclear medicine technologists. Part 1: Basic Principles of X-Ray Production. *J. Nucl. Med. Technol.*, **32**.

Seifarth, H., S. Wienbeck, M. Pusken, K. Juergens, D. Maintz, C. Vahlhaus, W. Heindel, and R. Fischbach (2007). Optimal systolic and diastolic reconstruction windows for coronary CT angiography using dual-source CT. *Am. J. Roentgen*, **189**.

Shemesh, J., S. Apter, J. Rozeman, A. Lusky, S. Rath, Y. Itzchak, and M. Motro (1995). Calcification of coronary arteries: detection and quantification with double-helix CT. *Radiology*, **197**.

Shope, T. B., R. M. Gagne, and G. C. Johnson (1981). A method for describing the doses delivered by transmission x-ray computed tomography. *Medical Physics*, **8**, pp. 488–495.

Silva, A. C., H. J. Lawder, A. Hara, J. Kujak, and W. Pavlicek (2010). Innovations in CT dose reduction strategy: Application of the adaptive statistical iterative reconstruction algorithm. *AJR, Am. J. Roentgenol*, **194**, pp. 191–199.

Taguchi, K., Z. Sun, P. W. Segars, E. K. Fishman, and B. M. W. Tsui (2007). Image-domain motion compensated time resolved 4D cardiac CT. *Proc. of SPIE.*, **6510**.

Tang, Q., J. Cammin, S. Srivastava, and K. Taguchi (2009). A fully four-dimensional, iterative motion estimation and compensation method for cardiac CT. *Med. Phys.*, **39**.

Tseng, H., J. Fan, and M. A. Kupinski (2015). Combination of detection and estimation tasks using channelized scanning linear observer for CT imaging systems. *Proc. of SPIE*, **9416**.

Tseng, H., J. Fan, M. A. Kupinski, P. Sainath, and J. Hsieh (2014). Assessing image quality and dose reduction of a new x-ray computed tomography iterative reconstruction algorithm using model observers. *Med. Phys.*, **41**.

Wagner, R. F., H. P. Chan, J. T. Mossoba, B. Sahiner, and N. Petrick (1998). Components of variance in ROC analysis of CADx classifier performance. *Proc. SPIE*, **3338**.

Wagner, R. F., H. P. Chan, B. Sahiner, N. Petrick, and J. T. Mossoba (1997). Finite-sample effects and resampling plans: Application to linear classifiers and computer-aided diagnosis. *Proc. SPIE*, **3034**.

Whitaker, M. K., E. Clarkson, and H. H. Barrett (2008). Estimating random signal parameters from noisy images with nuisance parameters: Linear and scanning-linear methods. *Optics Express*, **16**.

Wunderlich, A. and F. Noo (2008). Image covariance and lesion detectability in direct fan-beam x-ray computed tomography. *Phys. Med. Biol.*, **53**.

Young, S., P. R. Bakic, K. J. Myers, R. J. Jennings, and S. Park (2013). A virtual trial framework for quantifying the detectability of masses in breast tomosynthesis projection data. *Med. Phys.*, **40**.

Yu, L., S. Leng, L. Chen, J. M. Kofler, R. E. Carter, and C. H. McCollough (2013). Prediction of human observer performance in a 2-alternative forced choice low-contrast detection task using channelized Hotelling observer: Impact of radiation dose and reconstruction algorithms. *Med. Phys.*, **40**.