Performance optimization of a PET insert for simultaneous breast PET/MR imaging

PERFORMANCE OPTIMIZATION OF A PET INSERT FOR SIMULTANEOUS BREAST PET/MR IMAGING

BY

YICHENG LIANG, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF MEDICAL PHYSICS & APPLIED RADIATION SCIENCE

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF SCIENCE

Master of Science (2013) McMaster University

(Medical Physics & Applied Radiation Science) Hamilton, Ontario, Canada

TITLE:	Performance optimization of a PET insert for simultaneous breast PET/MR imaging
AUTHOR:	Yicheng Liang B.Sc., (Applied Physics) University of Waterloo, Waterloo, Canada
SUPERVISOR:	Dr. Hao Peng
NUMBER OF PAGES:	xi, 110

Abstract

Our group aims to develop a dedicated PET/MR system for breast cancer imaging. In order to evaluate and optimize the performance of the PET component, Monte Carlo simulation was made to help us choose the configuration parameters for hardware design. A resolution modeling method was also proposed and implemented on the GPU device to not only improve the spatial resolution of the reconstructed images but also accelerate the reconstruction speed.

The PET component is configured with a ring shape composed of LYSO+SiPM detectors. Such design is compatible to MRI, and feasible for time of flight PET. Several aspects are included to be investigated in the simulation which are geometry configuration, counting performance and image quality. From the simulation result, the system configured using $2 \times 2 \times 20mm^3$ LYSO crystal with two DOI layers and 3 detector rings results in 6.2% photon sensitivity. The Noise equivalent count rate is improved with better time resolution, the peak NEC is found to be 7886 cps with 250 ps time resolution. The system is able to achieve 2.0 mm spatial resolution which is found to be more uniform with the addition of DOI layers. With the help of TOF, the lesion is visualizable with shorter scan time than the non-TOF system.

The resolution modeling method is based on the coincidence detection response function modeling and multiray projection. It is found to improve the spatial resolution uniformity and contrast recovery. At the same time it reduces the position offset and background noise. The speed and accuracy improvement for this model is also discussed.

Acknowledgments

I would like to express my sincere appreciation to all those who have offered me invaluable help during the two years of my Master's study. Without your help, this thesis would not have been completed.

My deepest gratitude goes first and foremost to my supervisor, Dr. Hao Peng, who offered me the opportunity to study at McMaster University and gave me enormous help to my research. I would like to express my appreciation for your kindness, patience and constant encouragement. I am very proud to have such a nice supervisor.

I would also like give my hearty thanks to Dr. Troy Farncombe, Dr. Mike Patterson and Dr. Thomas Farrell, who kindly provide advice on my this project.

I must also thank all my friends, Farshad, Ibrahim, Elstan, Peter, Zahra, I will not forget about the great time spent with you. Thank you for all your kind help in campus living and course study.

I of course would like to thank my workmates, Xin, Ruibin, Kiara, Evan, David, Jenna. Thank you for the useful discussion.

Finally, I want to give particular thank to my parents, who give me all kinds of support, I love you all the time.

Notation and abbreviations

CNR	Contrast to noise ratio
DOI	Depth of interaction
FBP	Filtered backprojection
FOV	Field of view
FWHM	Full width half maximum
GPU	Graphic processing unit
LYSO	Cerium-dopped lutetium-yttrium oxyorthosilicate
LOR	Line of response
MLEM	Maximum likelihood expectation maximization
NEC	Noise equivalent count rate
OSEM	Ordered subset emission tomography
PET	Positron emission tomography
ROI	Region of interest
SIMD	Single instruction multiple data
SiPM	Silicon Photo-multiplier
SNR	Signal to noise ratio
SSRB	Single slice rebining
SUV	Standard uptake value
TOF	Time of flight
TOR	Tube of response

Contents

List of Figures

Chapter 1

Introduction and Background

1.1 Breast Cancer Imaging

Breast Cancer is one of the most frequent cancer among women with high mortality. Like many other types of cancers, the main strategy to manage the breast cancer is to detect the existence of lesions and plan the treatment in the early stage, with the use of medical imaging techniques. The widely used clinical breast imaging techniques include: X-ray mammography, ultrasound, breast MRI and breast PET.

The X-ray mammography is currently the most widely applied imaging technique for breast cancer screening. During the imaging procedure, the patients breast is compressed by two parallel plates to reduce the radiation dose, scattering and motion. Low-energy X-ray (30-60 keV) penetrates through the breast tissue and produces images on the screen film or digital devices placed on the opposite side. Breast cancer tissues typically have larger density than that of normal breast tissues due to micro-calcification. However, for patients who have dense breast tissues, there will be no contrast (i.e., morphological appearance) available between cancer tissues and normal tissues. Though the X-ray mammography is currently accepted as the golden standard for breast cancer screening, it

has relatively low sensitivity and specificity 55%-94% and 88%-99% (Khalkhali *et al.*, 1994), resulting in false positive diagnosis results which causes unnecessary biopsies, surgery and over-treatment.

Ultrasound is also being used for breast cancer patients, especially those who have ambiguous diagnosis with mammography. Instead of using radiation, ultrasound devices send high frequency sound waves through the breast tissue and measure echoes whose amplitudes are related to acoustic impedances of various tissues. It could also detect the blood flow(Sohn, 1992), and the pattern of vascularity(Delorme and Knopp, 1998) with the Doppler probe to increase the diagnostic accuracy, since the cancer tissues usually have more active blood flow and increased vascularity compared to benign tissues. In terms of sensitivity and specificity, the Ultrasound is reported to achieve 60% and 93% for breast cancer patients, respectively(Warner *et al.*, 2001).

Breast MRI tests are performed as supplemental tool for Mammography and ultrasound. It consists of a magnet with high magnetic field around patient body, three gradient coil sets and a surface RF coil around the breast. Besides relying on the contrast from the proton density (i.e., hydrogen atoms), contrast agents such as gadolinium are injected to enhance the contrast between lesions and normal tissues, based on the fact that the agents are more likely to accumulate inside regions nearby malignant tumors. To determine whether a lesion is benign or malignant, kinetic analysis could also be performed. However, since the basic diagnosis is still based on morphological differences (instead of physiological information), several limitations exist when it is used in detecting micro-calcification, inflammation or scar created by the biopsy or surgery(Morris, 2002). In terms of sensitivity and specificity, the MRI is reported to achieve 93% and 65% for breast cancer patients, respectively(Fischer *et al.*, 1999).

Positron Emission Tomography (PET) is another imaging modality, which is widely used in the staging and checking the re-occurrence of tumors after treatment(Gambhir, 2002).PET uses the distribution of radioactive tracers injected to access tumors at the molecular level, due to its high molecular sensitivity (nano-molar). The radioactive tracer inside body accumulates in the biological

regions, decays (beta plus) and emits positrons. Every positrons then annihilates with one electrons and create two 511 keV high-energy photons in the nearly back-to-back directions. The two 511 keV photons are then detected by the detectors around the patient in a narrow coincidence time window (∼several nanosecond). Images are produced after collecting millions of the high-energy photons pairs. Despite superior molecular sensitivity, a major limitation of PET in breast cancer patients is its limited spatial resolution (4-8 mm)(Doll *et al.*, 1997), which is not sufficient for detecting smaller lesions. In terms of overall sensitivity and specificity, the PET is reported to achieve 90% and 92% for breast cancer patients, respectively(Wu and Gambhir, 2003).

1.2 Breast dedicated PET system

To overcome the limit using the whole-body PET for breast imaging, breast-dedicated PET scanners have been proposed and developed in the last decade. Three major goals of the breast dedicated PET system instrumentation include: to increase the detection efficiency of 511 keV photons with optimized system geometry, to improve the ability of detecting small tumors using high resolution PET detector design, and to minimize the manufacture cost.

1.2.1 Coincidence events and signal quality in PET

When the two detectors collect two 511 keV photons in a fixed time window, it creates a coincidence event. Ideally, the event originates from a positron-electron annihilation and the tumor is assumed to stay along the line connecting two detector units, also called a line-of-response (LOR). This type of coincidence event is called true event. However, there are two other types of coincidence events in PET, both of which are not desired as illustrated in figure 1.1. The first one is scatter event. When one photon scatters, it will change direction and the two photons to be detected are no longer in the opposite direction. The second one is random event, when the two photons are resulted from two separate annihilations. As shown in figure 1.1, both scatter event and random event could create inaccurate LORs, which will degrade both spatial resolution and contrast recovery of lesions to be detected.

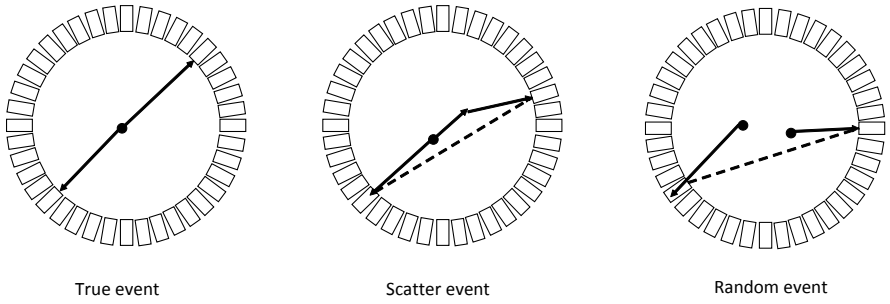

Figure 1.1: True, Scatter and Random coincidence events

Noise equivalent count rate (NEC or NECR), a good indicator of signal quality for a PET system, has been widely used to evaluate and optimize system performance. It is defined as below:

$$NEC = \frac{T^2}{T + S + R} \tag{1.1}$$

where T, S and R represents the true, scatter and random coincidence count rate. From the formula above, there exist two ways to improve the NEC: 1) increase the total counts following Poisson statistics; 2) minimize scatter and random events.

Photon detection efficiency

The significance of improving detection efficiency is to ensure the system to get adequate counts and signal-to-noise (SNR) per pixel in the image. In addition, high detection efficiency could also reduce the activity of the radio-tracer injected, as well as the acquisition time and radiation dose exposure. Increasing the photon detection efficiency (i.e. geometric efficiency) could be achieved by placing the detector closer to the objects (i.e., breast), resulting in larger solid angle coverage. Meanwhile , such design also requires less scintillation crystals and helps reduce the cost. Several groups have proposed designs of breast dedicated scanner in particular shapes, including dual-panel(Abreu

et al., 2007)(Zhang *et al.*, 2007)(MacDonald *et al.*, 2009)(Peng and Levin, 2010), rotating dual-panel(Camarda *et al.*, 2006)(Raylman *et al.*, 2008)(Wu *et al.*, 2009), rectangular(Wang *et al.*, 2006), complete ring(Moliner *et al.*, 2012) and partial ring(Furuta *et al.*, 2009a)(Surti and Karp, 2007).

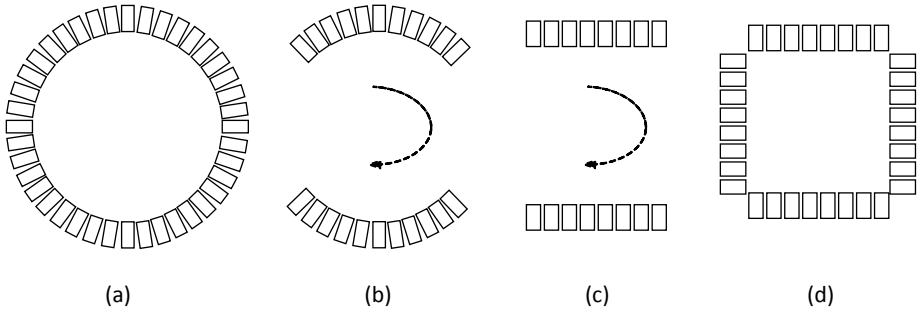

Figure 1.2: Different geometry configuration of breast dedicated PET systems. (a) Full ring. (b) Partial ring. (c) Dual-panel. (d) Rectangular. (b) and (c) could be either static or rotational.

The photon detection efficiency could also be improved by increasing the detection efficiency of individual block detectors (i.e., intrinsic efficiency). Typically, a PET detector is composed of scintillation crystals, photodetectors and electronic readout systems. The scintillation crystal converts the 511 keV photon to visible lights which travel to the coupled photodetectors and are then converted to electronic signals. By using scintillation material of high atomic number (Z) and/or increasing the thickness of scintillator materials, we can enhance the photon detection efficiency. The size of the scintillation crystal could also alter the detection efficiency. When the crystal is too small, larger portion of 511 keV photons will penetrate through the crystal without interacting with it.

Energy and time resolution settings

Once the electronic system detects the signals from two photodetectors, the following signal processing is performed either in hardware or software mode. First, when a high-energy photon undergoes Compton scattering, a portion of its energy will be transferred to the Compton electron, resulting in a smaller amount of energy to be deposited within the detector. These scatter events can be

rejected by setting up an energy window (typically 20-30% of 511 keV on either side of the 511 keV photopeak). On the other hand, since those random events are more likely to have large time difference, a small time window (typically 4-6 ns in a whole body PET system, and about 2-3 times the time resolution) can be applied to differentiate them from true coincidence events. In brief, under a given photon detection efficiency, a PET system of superior energy resolution and time resolution is expected to obtain higher NEC and improved performance.

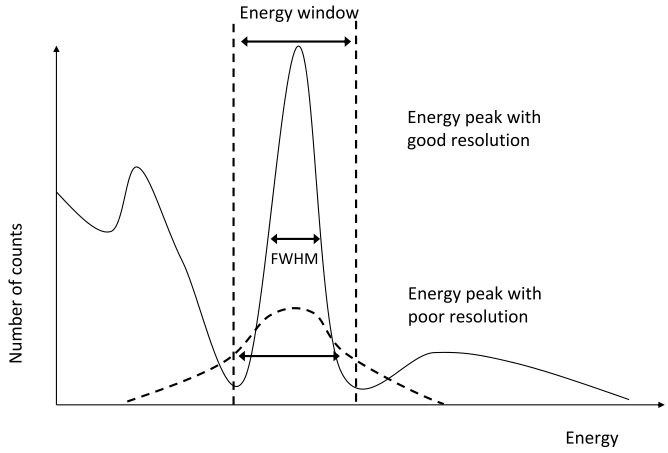

Figure 1.3: Energy resolution and energy window. Energy peak with smaller FWHM reflects better energy resolution of the detector.

1.2.2 PET spatial resolution

Another important motivation of the breast dedicated PET scanner is to detect small tumors where the standard whole body scanner can not. The ability of distinguishing two adjacent small points in an image is referred as the spatial resolution. For a PET system, this is determined by several factors: detector configuration, positron annihilation physics, depth-of-interaction effect, image reconstruction and processing. A semi-empirical framework has been developed to predict the spatial resolution of a PET system based on measurements (Sorenson *et al.*, 1987) as shown

below.

$$R_{system} \approx \sqrt{R_{det}^2 + R_{range}^2 + R_{180^o}^2} \tag{1.2}$$

The overall spatial resolution associated with detector response, positron range, photon acollinearity, as well as the image reconstruction process, which will be discussed in Chapter 3.

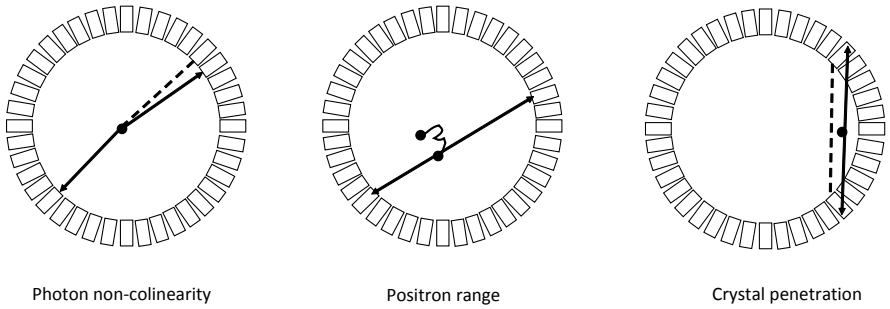

Figure 1.4: Illustration of three blurring factors for PET resolution: photon non-collinearity, positron range and crystal penetration

The first term (crystal pitch) is one dominant factor, to be determined by the pitch of the scintillation crystal, inside which no accurate positioning information is to be known. Whole body PET scanners deploy crystals of 4-6 mm pitch, while for most breast dedicated scanners, the crystal pitch is reduced down to 1-2 mm for achieving better spatial resolution.

The second term (positron range) originates from the physics of positron-electron annihilations. When a positron is created, it carries certain amount of kinetic energy and travels in a distance before its annihilation with an electron. For a given radionuclide, the positron range is dependent on the energy distribution of beta plus emission(Phelps *et al.*, 1975).

The third term (non-collinear annihilation) also originates from the physics of positron-electron annihilations. When a positron annihilate with an electron, they will not create a pair of photons in the exact 180-degree direction, due to the residual momentum of the positron. The effect of

this acollinearity on the spatial resolution is proportional to the diameter of a PET scanner (cylindrical configuration), or the separation of two detector units for other non-cylindrical configurations.

Another resolution blurring factor not explicitly included in the formula is depth-of-interaction (DoI), also known as parallax error. Towards the edge of field-of-view (FoV) of the system, the spatial resolution degrades through the crystal penetration effect as illustrated in figure 1.4. For example, if a breast PET system is to be developed right around the breast (15-18 cm diameter) and 20 mm thick crystals are to be used, the resolution is expected to degrade significantly from the center to the edge. With the DoI effect, the detector response in terms of FWHM is expressed as $R_{det} \approx d/2 \times [cos\theta + (x/d)sin\theta]$ where d is the crystal width, x is the crystal thickness. One can mitigate the DoI blurring by utilizing more advanced crystal array design.

1.2.3 PET basics and breast dedicated PET

Table 1.1 List of breast dedicated PET systems

scanner	geometry configuration	crystal size(mm^3)	spatial resolution	sensitivity
clear-PEM	2 rotating panels	2 × 2×20 LYSO	2.5 mm	4.3%
Flex Solo II	2 panels	2 × 2×13 LYSO	2.4 mm	18%
MDA-PEM	ring	2.7 × 2.7×18 BGO	2.7 mm	9.2%
PEM/PET	2 rotating panels	2 × 2×15 LYSO	2 mm	6.88%
LBNL PEM	rectangular	3 × 3×30 LSO	2.1 mm	4.9%
MAMMI	ring	monolithic LYSO	1.9 mm	1.8%
C-PEM	C-shape partial ring	1.44 × 1.44×4.5 LGSO	1.2 mm	6.9%
PEM-CT	2 rotating panel	3 × 3×20 LYSO	2.5 mm	1.64%
CZT	2 panels	5 × 40×40 (CZT)	1 mm	32%

A number of breast dedicated PET system designs are summarized in Table 1.1 and briefly discussed here. It is noticed that the flex solo II scanner designed by Naviscan PET systems has been approved by FDA to clinical application, which has integrated the capability for biopsy guidance (Kalinyak

et al., 2011). Other systems are still in the phase of prototype and lab research. When comparing these dedicated systems to a standard whole body PET system, there exist several major differences.

First, all of these systems use smaller scintillation crystal sizes (∼2 mm) than the whole body system (4-6 mm) and is thus able to achieve improved spatial resolution. Second, the dedicated systems are placed very close to the breast (either panel or ring geometry), which allows it to achieve higher photon sensitivity. It is noticed that the stationary dual-panel systems have much higher sensitivity due to its smaller panel separation (∼4cm) to compress the breast during the acquisition. Third, the cost of building a breast dedicated PET scanner is less than that of the whole body scanner. The cost of a whole body system is around $1.5 million US dollars (Von Schulthess *et al.*, 2000). In contrast, the dedicated systems are expected to be more cost-effective (10%∼20% of the whole body system) since the reduced number of scintillation crystals and detector are required. Fourth, the compact size of the dedicated systems requires less space and makes the systems more portable.

The fifth aspect is about how to improve resolution uniformity, which is a major challenge to be addressed by these dedicated systems. The potential resolution non-uniformity results from either parallax error due to reduced system diameter or panel separation, or limited angular coverage (Budinger, 1998). As shown in Table 1.1, several systems have incorporated depth of interaction (DoI) capacity to reduce the parallax error(Abreu *et al.*, 2007)(Wang *et al.*, 2006)(Furuta *et al.*, 2009a)(Moliner *et al.*, 2012)(Peng and Levin, 2010). For those systems that has limit angular coverage, the geometry configuration increases the flexibility to place the scanner around the breast to image the whole breast including the area that is close to chest wall and axilla. For example,(Furuta *et al.*, 2009b) designed a C shaped scanner which could place the entire breast in the FOV. However, the limited angle coverage design causes resolution blurring and/or image artifacts due to missing sampling at some projection angles. There are three possible solutions to this challenge in the recent literatures. The first one is to rotate the two panels in order to acquire data in all

directions(Camarda *et al.*, 2006)(Raylman *et al.*, 2008). The second one is to incorporate the time-of-flight technique (to be discussed in details in section 1.4) using new scintillation crystal such as $LaBr_3$ and LYSO (Surti and Karp, 2008). The third one is to further improve DoI capability with advanced PET detector designs to help mitigate resolution blurring along the direction perpendicular to the two panels. For example, CZT detectors have been tested to replace the scintillation crystal. The high atomic number and superior energy resolution at 511 keV (3%) result in high detection efficiency for a PET system. In addition, the cross-strip readout scheme of CZT could provide DoI capability down to 1 mm with electronically pixelated detector units(Peng and Levin, 2010), which could not be easily accomplished in cutting the scintillation crystals into small pitches.

1.3 Breast PET/MRI

Unfortunately, a single imaging modality usually can not meet all the clinical requirements. Multi-modality imaging devices are then developed to utilize more than one of the imaging methods to provide more information. For example, the current clinical PET systems are usually combined with the CT device, in order to obtain the high resolution anatomical information and sensitive functional information at the same time. Another benefit of the PET/CT is that the CT data could be used for the attenuation correction, which improves the image quality for the PET image. People have also tried to combine MRI and PET to build PET/MR devices, as compared to CT, MRI is expected to provide better soft-tissue contrast and result in no radiation dose.

Currently, there are three design concepts for PET and MRI combination(Werner *et al.*, 2012). The first one is separate the PET and MRI scan in different rooms. The second is the co-planar design where patient is lying in the same plane for PET and MRI in the same room. The last one is the fully integrated system where PET and MRI scanned simultaneously. The first two designs scan PET and MRI image separately, where patient motion may create difficulty in the image fusion. The fully integrated system allows both component to scan simultaneously, therefore, image blurring caused by motion could be corrected, in addition, the total scan time could be significantly reduced.

However, the design of fully integrated simultaneous PET/MR has many challenges(Sauter *et al.*, 2010). First, the integration of the PET component could result in limit space for imaging, the bulky PET component such as PMT should be replaced to more compact component. In addition, the strong magnetic field could affect the PET detectors especially the PMT, where the electrons path between the dynodes in the PMT could be altered. Furthermore, the switching of the gradient field could create eddy current on the PET component which could affect magnetic field and result in artifact in the MR image. Finally, compared to PET/CT, the attenuation correction for PET/MR is a great challenge as MR uses a complete different mechanism as for CT, no x-ray will penetrate through the body.

Several research groups are developing methods for fully integration of PET/MR. The first solution is to place the scintillation crystal inside the MRI scanner and place the photo-detectors such as PMT, APD away from the crystals connected by optical fibers (Shao *et al.*, 1997)(Catana *et al.*, 2006). However the drawback for this design is that the scintillation light is significantly dispersed and attenuated which could consequently degrade the detection efficiency and image quality. The second solution is to use solid state photo detectors such as APD and SiPM to build PET insert which could be placed inside a MRI scanner (Catana *et al.*, 2008)(Judenhofer *et al.*, 2008)(Judenhofer *et al.*, 2006) without optical fibers connecting scintillation light outside of the MRI component. To reduce the interference with the magnetic field, the charge sensitive preamplifier electronics are placed inside a thin copper shielding. However, in order to fit in the standard MRI scanner, usually the size of the PET insert is limited which is used for small animal imaging study. There remains challenges for a fully integration of PET and MRI for a clinical whole body PET/MRI scanner.

In terms of breast imaging, there are three modes to combine PET and MRI. The first mode is operated in a mode of prone breast MR and supine PET, in which two image sets are evaluated either through visual fusion or software fusion. In both cases, a serious challenge is the different

position of patient in the two separate scans and the co-registration of small lesions. The second mode is in a mode of prone breast MR and prone PET. Prone PET was achieved by having the patient remain in a position identical to the MR using a conventional MR breast holder (but with the receiver coil removed). Clinical studies that fused PET and MR scans increased the specificity but decreased the sensitivity, and additional data are needed to confirm the significance of image fusion(Goerres *et al.*, 2003)(Moy *et al.*, 2007). In addition, the studies discussed two factors that limit the accuracy of co-registration: 1) position and shape changes observed for nearly all patients (especially for women with large breasts); and 2) the size limitation of a lesion that can be detected by a whole body PET (4-6 mm) poses a challenge on quantitative analysis of specific uptake value. The third mode of fusion is truly simultaneous prone PET/MR. It is expected that a simultaneous PET/MR acquisition is desired over two separate scans, by facilitating image registration, patient handling, and reduces scan time. Besides, precise temporal correlation of the complementary signals will lead to a powerful tool allowing us to exploit the full potential of MRI such as diffusion-weighted imaging and MR spectroscopy(Buchbender *et al.*, 2012). A prototype breast PET/MR has been investigated at Brookhaven National Lab prototype based on RatCAP small animal scanner(Ravindranath *et al.*, 2009).

1.4 TOF PET and silicon photomultiplier (SiPM)

1.4.1 TOF

Besides the PET principles discussed in section 1.2, there are two recent developments in PET detector/system highly relevant to this thesis: TOF PET and SiPM.

The TOF PET measures the time difference of the two 511 keV photons to localize the position of the annihilation in a smaller region. Compared to the conventional systems, the TOF PET systems can reduce the noise in the reconstructed image significantly (Budinger, 1983)(Karp *et al.*,

2008), which is due to several reasons: firstly, for the conventional systems, the noise from all the pixels along a line-of-response (LOR) is correlated; secondly, in a TOF PET system, the statistical fluctuations from the data along a LOR are constrained to a reduced number of image pixels, as illustrated in Figure. 1.5.

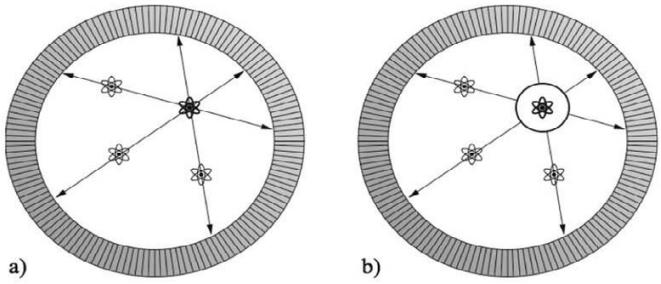

Figure 1.5: (a)In conventional PET, the source is localized to a LOR between a detector.(b)In TOF-PET, timing information is used to restrain the source to a segment of the LOR

$$SNR_{TOF} = \sqrt{\frac{D}{\Delta x}} SNR_{non-TOF} \tag{1.3a}$$

$$\Delta x = \frac{c}{2}\Delta t \tag{1.3b}$$

The SNR benefit of a TOF PET is given by equation (1.3 a)(Conti, 2009), where Δx is the position uncertainty defined by equation (1.3 b), c is the speed of light, Δt is the time difference, and D is the diameter of the scanner. Back to the time the TOF concept was introduced in the 1980s, PET systems were mainly used for brain and cardiac imaging whereas resolving small lesions and improving SNR are not considered very critical. However, the TOF PET has gained more research and clinical attention as PET increases its role in oncology nowadays. Several simulation and experimental studies have been performed to predict the performance of whole-body TOF PET systems. Phantom studies with a LYSO-based commercial TOF PET scanner (Philips Gemini TF PET/CT) indicate that TOF information not only increases the lesion detection ability by a SNR gain of 1.4-1.6 (10 mm diameter lesion within 35 cm diameter background), but also reduces the

number of iterations required in statistical image reconstruction(Karp *et al.*, 2008). Clinical studies also provide promising results that a TOF PET is able to reveal/define structural details in colon cancer, abdominal cancer and lymphoma(Karp *et al.*, 2008). Currently, a number of TOF PET scanners are commercially available.

1.4.2 SiPM

The SiPM is a novel photo-detector for the TOF PET system development(Herbert *et al.*, 2007)(Spanoudaki and Levin, 2010). The SiPM detector consists of many SiPM pixels ranging from 1 mm to 3mm, Each pixel contains lots of micro-cells. Each micro-cell of 20-50 μm pitch operates in Geiger mode, which is also called Geiger APD and single photon avalanche diode (SPAD). The typical gain of a micro-cell is 10^5-10^7, which is very comparable to that obtained with PMTs. Compared to PMTs, SiPM offers the solid-state advantages such as low operating voltages, ruggedness, compactness and excellent immunity to magnetic fields.

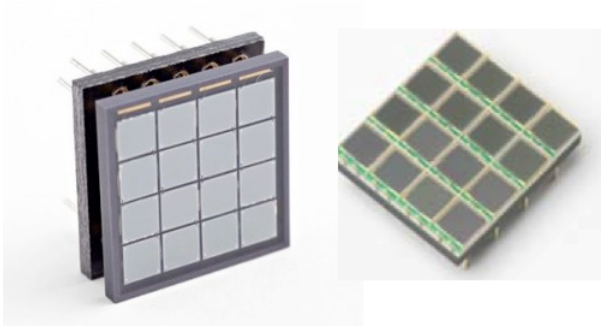

Figure 1.6: 4 × 4 array of SiPM with 3.0 mm pitch from SensL (left) and Hamamatsu (right)

The SiPM detector combines a number of advantages for PET instrumentation. First, its small pixel size (1-3 mm) enables detectors to achieve high spatial resolution. Second, its high gain results in good detector SNR and eliminates the need for subsequent stages of amplification. Third, the output pulse of each micro-cell has a well-defined pulse shape (i.e. rising edge and amplitude),

which is advantageous for accurate timing pick-off(Spanoudaki and Levin, 2010). SiPMs have been used to design PET detectors. In particular, recent studies show that a time resolution of 237 ps FWHM and 240 ps FWHM is achieved for LaBr3(Schaart *et al.*, 2008) and LYSO(Kim *et al.*, 2009) crystals of 3x3x20 mm^3 pitch, respectively.

1.5 Motivation

Our group aims to develop a PET/MR system for breast cancer imaging, which comprises scintillation crystals and SiPM arrays and is able to be integrated with MR breast coil to perform simultaneous PET/MR imaging. Our purposed PET ring insert is expected to have the following features: 1) a reduced system diameter (15 cm) to locate right around the breast for higher photon detection efficiency, 2) high spatial resolution (1-2 mm) for detecting small size lesion for early detection; 3) compatible with a breast MR; 4) lower cost.

To optimize the design of the PET component of this system, Monte Carlo simulation is performed in order to help us choose the configuration parameters and evaluate the performance before building and testing a real system, which is going to be covered in the chapter 2. In the third chapter, a GPU accelerated resolution modeling method is proposed and studied in order to improve the image quality from the image reconstruction process. Chapter 4 is about the conclusion and future work, followed by the appendix which lists image reconstruction and correction methods.

Chapter 2

Performance optimization using Monte Carlo simulation

This chapter focuses on the performance optimization of the proposed PET insert using Monte Carlo simulation. Several aspects are included to be investigated in the simulation. The first aspect is choosing the size and number of scintillation crystals, in order to achieve $1 \sim 2$ mm spatial resolution, 1mm and 2mm scintillation crystals are chosen as candidates. The number of scintillation crystals that forms a detector ring and the number of rings are to be selected to make sure the system covers sufficient FOV, maintain enough counting efficiency, and minimize the cost. The second aspect is the counting performance of this system, which focuses on reducing the random and scatter events while improving the true events. Two methods are going to be tested, the first one is choosing a detector system with optimized time window and energy window setting, and the second one is placing the shielding material around the device to reject unexpected events from the body. The last aspect is the image quality, which could be influenced by both the system design and post processing by the image reconstruction program. The image quality is tested by the spatial resolution and contrast recovery, to improve them, techniques such as time of flight and resolution modeling is to be incorporated and evaluated.

2.1 Monte Carlo simulation with GATE

The Monte Carlo simulation, is a method that uses mathematical model based on random numbers to solve physical problems. The Monte Carlo simulation has a wide range of applications. In the field of PET, it is mainly applied to optimize the system design, evaluate image reconstruction methods and model radioactive tracer kinetic (Buvat *et al.*, 2002). The key simulation model was built from the random characteristic of radioactive decay, photon interaction and electron interaction. A number of softwares have been developed to model the physical processes (Rogers, 2006), such as EGS4 (Nelson *et al.*, 1985), MCNP (Briesmeister, 1986), GEANT4 (Agostinelli *et al.*, 2003) and PENELOPE (Baro *et al.*, 1995). Based on the Monte Carlo codes for photon physics, dedicated softwares for PET were also built, including GATE (Jan *et al.*, 2004), PETSIM (Thompson *et al.*, 1992), PETEGS (Castiglioni *et al.*, 1999). This work uses the GATE simulation toolkit to study and optimize the performance of a breast dedicated PET ring insert for PET/MRI imaging.

GATE is developed by the openGATE collaboration, it is an opensource software based on GEANT4 designed for emission tomography. The simulation is coded using macro scripts instead of writing specific c++ programs. The whole simulation processes could be summarized as three processes: system geometry configuration, physical models setting, and data digitization and output.

The system geometry configuration of the whole simulation includes setting up the PET scanner and the phantom. The PET scanner is usually composed of one or more detector rings, which is composed of many block detectors, and each block detectors is made up of scintillation crystals. By defining the material and size of the scintillation crystals and repeating, detector blocks are constructed. And by repeating the detector blocks, ring detectors and the whole PET scanner are made. Phantoms are also placed within the scanner for acquisition. The phantom could be a simple cylinder or a complex anthropomorphic phantom used for performance evaluation. The phantom could also be simulated with motion, allowing people to study the effect of motion on the

image. GATE inherits the physics packages in GEANT4 which are well validated for PET simulation. It also allows users to define the source type, geometry, activity, decay time and motion. The digitizer function allows users to simulate the detector readout systems to select and record coincidence events through a coincidence chain. The following section will cover more details about the simulation process to study and optimize the performance of a breast dedicated PET insert for simultaneous PET/MRI imaging.

2.2 System configuration

The configuration of the proposed PET insert is shown in Figure 2.1 The PET insert will be located very close to the breast, which enables large solid angle coverage and high photon detection sensitivity. The outer diameter of the ring is set to be 19 cm so that the insert can fit into the MR breast coil. In order to reduce physical space and be cost-effective, the insert is not designed to cover the full breast along the vertical direction and therefore the number of rings needs to be optimized to study the trade-off between system performance (sensitivity) and cost. The detector module is to be based upon SiPM detectors (magnetic field compatible) and LYSO (Cerium-doped Lutetium Oxyorthosilicate) crystals.

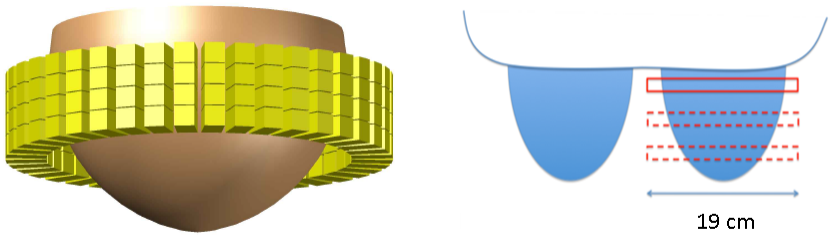

Figure 2.1: Sketch diagram (left) of the proposed breast PET/MR system and side view (right) of the PET ring around the breast (prone position), to be integrated with commercially available breast MR coils. The PET system can be translated vertically to cover the full breast.

Four system and crystal designs in our simulation are summarized in Table 2.1. The first two

designs (I and II) focus on the dependency of performance on two crystal pitches (1 mm vs. 2 mm), both of which are smaller than that of a standard clinical whole body PET scanner (4 mm). Designs III and IV focus on the dependency of spatial resolution across the field-of-view (FoV) on DOI capability, which is a potential challenge for small ring geometry. The two-layer DOI design is to be implemented through either dual-layer readout or triangular-rectangular crystal design(Peng and S Levin, 2010).

Table 2.1. Four system and crystal configurations studied in the simulation

	I	II	III	IV
Crystal dimension (mm^3)	1x1x20	2x2x20	1x1x20	2x2x20
DOI resolution (mm)	20	20	10	10
Crystals per block	64	25	128	50
Blocks per ring	46	40	46	40

Based on the experiment results in our group(Yang *et al.*, 2013)(Downie *et al.*, 2013), the LYSO and SiPM module has a time resolution of 1ns FWHM and an energy resolution of 12% FWHM centered on 511 keV. In the simulation, two opposite 511 keV photons are used instead of positron sources which annihilate with electrons in order to speed up the simulation. As the diameter of the scanner is about 5 times smaller than the whole body scanner (70cm), photon acollinearity effect is less dominate in the spatial resolution blurring. The positron range could also be modeled and applied to the image after the simulation (Levin and Hoffman, 1999). Therefore using two photons could speed up the simulation without changing the result significantly. In GATE, each interaction in the sensitive detector (phantom or crystal) is recorded as a "Hit", the Hits could be grouped together in the pre-defined volume unit and output to Singles file. The Singles file lists the ID of the annihilation event (eventID), the ID of the detected crystal (crystalID), the summed energy of all the Hits (energy), the time of the last Hit (time) and the number of Compton scatter occurred in the phantom (scatterID). These data could be used to sort coincidence events in the GATE by defining a coincidence processing chain and importing them in the chain. For this simulation, the

coincidence events are processed after the simulation using a in-house C program, where parameters such as time and energy window could be changed without running the GATE simulation again.

The data recorded in the singles files such as time and energy are the exact value. However, there is no perfect equipment to record such value, so the time is blurred by a Gaussian function defined by the time resolution, and the energy is blurred by another Gaussian function defined by the energy resolution. After the time values are blurred, the pre-defined time window is created, if only two Singles occur in the time window and the energy values are within the energy window, then they are considered as a pair of coincidence events. For the coincidence events, if the eventIDs for the two Singles are different that means they are created by two annihilation events, then they are considered as random events. If the eventIDs are identical but the scatterID is greater than zero, the photon undergoes Compton scatter in the phantom, then they are considered as scatter events. If eventIDs are identical and scatterIDs are zero, then they are true events. The processed data are used for the system performance evaluation and image reconstruction.

2.3 Method

2.3.1 Photon sensitivity

The first merit for the system performance is the photon sensitivity. The photon sensitivity is defined by the ratio of the true count rate to the source activity. The larger the photon sensitivity, the higher the detection efficiency, and the shorter scan time or the lower the radiation dose. In this study, the sensitivity is determined by placing a point source with known activity in the scanner to measure the true count rate. To test the spatial dependence, the position of the point source is changed. The number of ring detectors is also adjusted in order to balance the trade off between photon sensitivity and cost.

2.3.2 Noise equivalent count rate

Another merit for the system performance is the Noise equivalent count rate (NEC) defined by equation 2.1 (Strother *et al.*, 1990).

$$NEC = \frac{T^2}{T + S + R} \tag{2.1}$$

where T, S, R are true, scatter and random coincidence events respectively. To study the impact of other organs adjacent to breast on the NEC performance, the XCAT anthropomorphic phantom is used to mimic the true attenuation and source distribution in the body. The XCAT phantom (Segars and Tsui, 2009) is a voxelized human phantom based on high resolution multislice CT data. It is composed of two separate files, one for attenuation distribution and the other for activity distribution, both files are loaded into GATE to construct the voxelized phantom (voxel size: $4 \times 4 \times 4$ mm^3). The activity distribution in different organs is listed in table 2.2. based on literatures(Ramos *et al.*, 2001) (MacDonald *et al.*, 2009).

Table 2.2. List of activity concentrations in various organs

Organ	Breast	Heart	Liver	Spleen	Bladder
(uCi/mL)	0.10	1.03	0.43	0.29	2.93

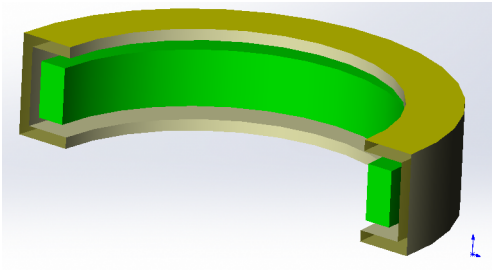

Figure 2.2: Lead shielding around the PET ring insert

The dependency of NEC on different energy window and time window settings were studied. Furthermore, in order to study the potential effect of TOF capability, a comparison of NEC performance

for different time resolutions ranging from 8ns to 250ps was performed. For each set of simulation, the number of coincidence events was approximately 400000. In addition, NEC was also studied as a function of injected activity concentration and lead shielding around the PET insert.

2.3.3 Spatial resolution

Two sets of simulations were made to evaluate the spatial resolution and DOI effect in this section. First, we constructed micro-Derenzo hot rod sources with the activity concentration set to be 1 μci/cm^3 in a cold background. Second, to study the in-plane spatial resolution as a function of radial offset, rod sources with a diameter of 0.4 mm were placed along radial lines from the center to the edge of the scanner (the distance between any two sources was set to be 1 cm). Filtered back projection (FBP) with single slice rebinning technique(Daube-Witherspoon and Muehllehner, 1987) was used for image reconstruction. Due to the large number of gaps in-between two adjacent detector blocks (i.e., small ring diameter), the regions of zero intensity in a sinogram were filled using an interpolation method as described in(de Jong *et al.*, 2003). A uniform cylindrical source was used for normalization correction in the image reconstruction. For the DOI study task, line profiles were obtained from the reconstructed images and analyzed quantitatively for FWHM values. Four system designs were studied.

2.3.4 Contrast analysis

To evaluate lesion detectability of the proposed system, four sphere lesions with different diameter values (5 mm, 4 mm, 3 mm and 2 mm) were placed in the left breast (1.5 cm away of chest wall of the XCAT phantom). Activity concentration in the lesions was chosen to be 10:1 and 5:1 with respect to the background (0.1μCi/cm^3). Images were reconstructed using the 3D list-mode OSEM method with 6 subsets and a set of iteration numbers for optimizing CNR. The voxel size was

selected to be $1 \times 1 \times 1\ mm^3$. Only design IV (2 mm crystal and 10 mm DOI) was chosen for contrast study.

To investigate any potential benefit of TOF for lesion detectability, three time resolution values (1 ns, 500 ps, and 250 ps) were studied. For the last two cases that reflect the present limit of achievable time resolution of a scintillator and SiPM detector, the images were reconstructed incorporating TOF information. For instance, the distance between source position and the center of the line of response x was calculated using Equation. 2.3a, where c represents the speed of light and t is the time difference between the two detectors. The TOF kernel of the line of response in the reconstruction was determined by the time resolution as also shown in Equation. 2.3c (Pratx *et al.*, 2011) where the spatial invariant Gaussian kernel (Equation. 2.3b) is now multiplied by the TOF Gaussian kernel and the standard deviation of the TOF kernel is determined by the time resolution using Equation. 2.3a.

$$\Delta x = \frac{c}{2}\Delta t \tag{2.2a}$$

$$K(d_{ij}) = exp(\frac{-d_{ij}^2}{2\sigma_x^2}) \tag{2.2b}$$

$$K(\Delta x) = K(d_{ij})exp(\frac{-\Delta x^2}{2\sigma_t^2}) \tag{2.2c}$$

$$CNR = \frac{c_{lession} - c_{background}}{\sigma_{background}} \tag{2.2d}$$

To analyze the CNR shown in Equation. 2.3d, regions of interest (ROIs) were chosen for quantifying activities in the lesions (C_{lesion}) and background ($C_{background}$) respectively. $\sigma_{background}$ stands for the standard deviation of the background(Gu *et al.*, 2010). The CNR was studied as a function of both imaging time and iteration numbers. The results were compared against Rose criterion as a semi-quantitative metric, which indicates that good lesion detectability can be obtained for a CNR value above 4(Rose, 1973). It should be pointed out that for the 2 mm lesion (image pixel size: 1 mm), it is quite challenging to identify the appropriate ROI and thus no quantitative analysis was made for the activity concentration ratio of 5:1.

2.4 Result

2.4.1 Photon sensitivity

First, the relationship between photon sensitivity and the number of detector rings is studied. A point source with 100 μCi is placed in the center FOV, the time and energy window are set to be twice of the time and energy resolution respectively. Since the introduction of DOI will not affect photon sensitivity, only design I and II were studied. For both cases, the photon sensitivity increases linearly as the number of detector rings increases, though associated with increased system cost. For example, for the design I (1 mm crystal size), 4 rings (3.3 cm axial height) result in a peak sensitivity of 4.6%; while for the design II (2 mm crystal size), 3 rings (3.3 cm axial height) result in a peak sensitivity of 6.2%. As the number of rings doubles (increasing system cost and space requirement), the peak sensitivity increases to 9.6% (design I) and 12.7% (design II), respectively.

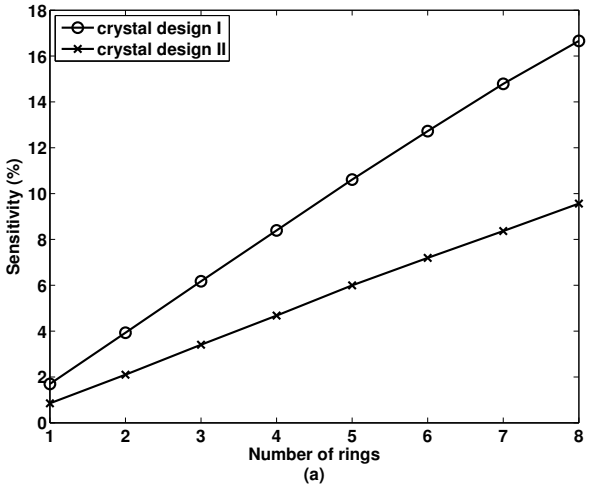

Figure 2.3: Detection sensitivity of 511 keV photons using a point source as a function of number of detector rings

Second, the photon sensitivity is tested by the position of the point source in the axial and transaxial directions as shown in figure. The sensitivity decreases nearly linearly from the center of FoV towards the edge in the axial direction. In the trans-axial direction, the sensitivity profile exhibits an oscillating pattern, mainly because of the gaps in-between crystal blocks and hence non-uniform detection efficiency. Besides, the sensitivity in the trans-axial direction increases as the point source shifts towards the edge of FoV, when the source is placed closer to the detectors and results in larger solid angle coverage.

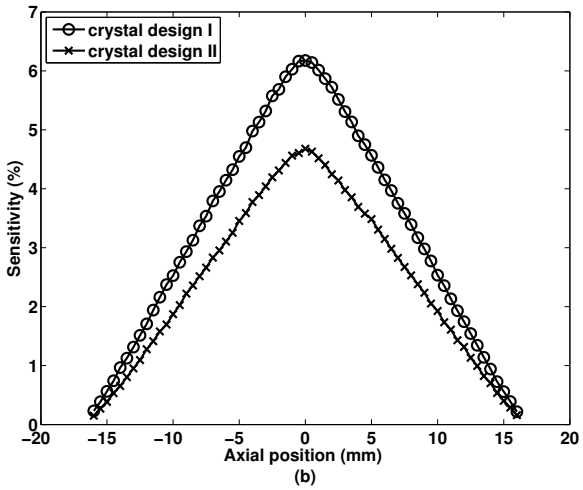

Figure 2.4: Detection sensitivity of 511 keV photons using a point source as a function of axial position

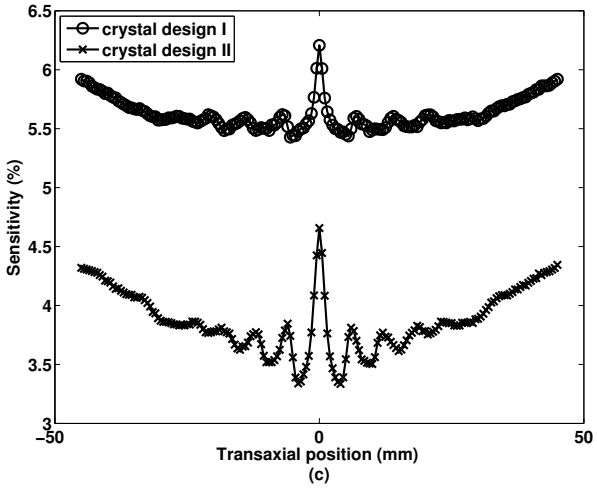

Figure 2.5: Detection sensitivity of 511 keV photons using a point source as a function of transaxial position

2.4.2 Noise equivalent count rate

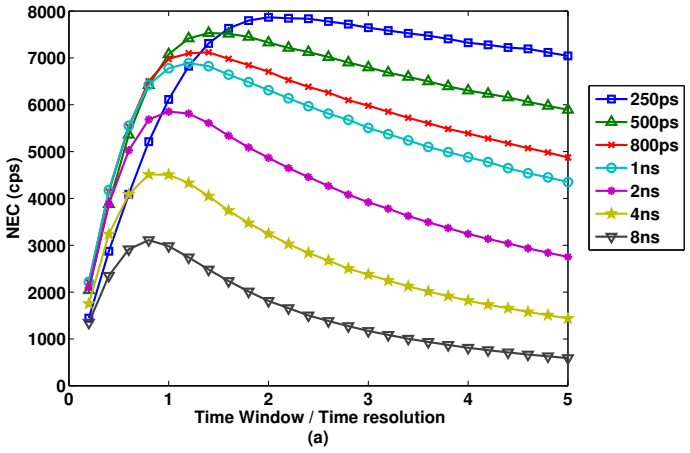

Figure 2.6: NEC result as a function of the ratio between time window and time resolution

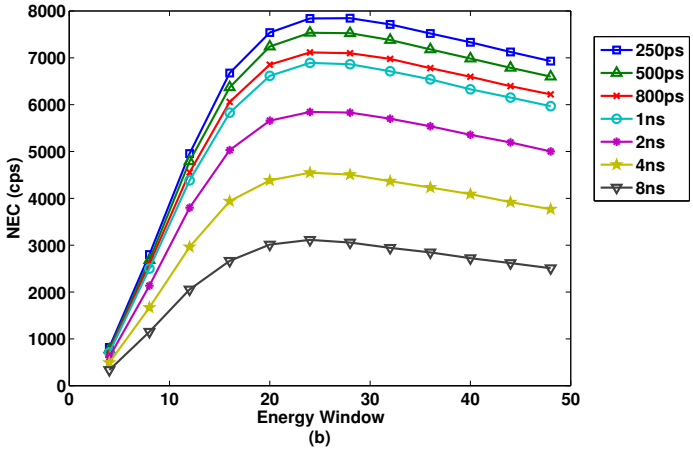

Figure 2.7: NEC result as a function of energy window centered at 511 keV

Figure 2.6 and Figure 2.7 show the NEC dependence on time and energy window for a set of time resolution settings. The energy resolution is 12% centered at 511 keV for all cases. The results indicate that for the proposed system, improved time resolution can increase the NEC by rejecting random events (as well as multiple coincidences) effectively. High count rate of random coincidences is a serious challenge associated with systems of small diameter PET systems. The highest NEC peak value is found to be 7886 cps for 250 ps time resolution with a time window of 500 ps. The lowest NEC peak value is found to be 3109 cps for 8 ns time resolution with a time window of 6.4 ns. Furthermore, it is observed that for different time resolutions, the peak NEC occurs at different time window settings. In other words, the ratio of time window to time resolution resulting in peak NEC ranges from 0.8 to 2.0. On the other hand, the energy window for achieving peak NEC is found to be 24% (twice the energy resolution) for all studies. As the width of energy window increases, more scatter events are included and thus result in a gradual decrease of NEC.

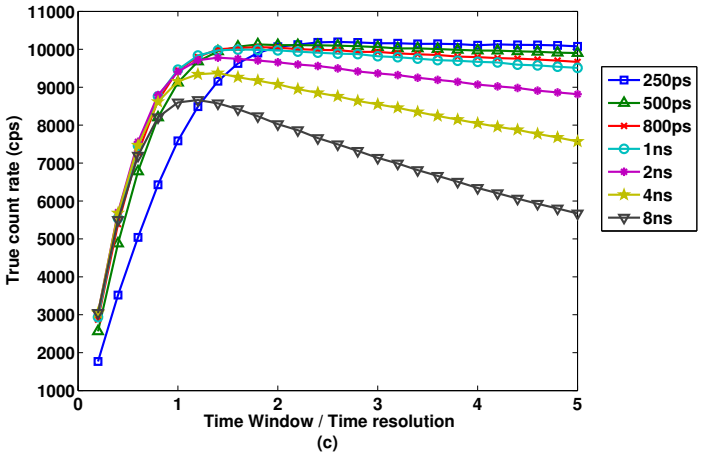

Figure 2.8: True event count rate as a function of the ratio between time window and time resolution

Figure 2.8 shows the true coincidence count rate as a function of ratio between time window and time resolution. It is noticed that as the ratio increases, the true count rate increases initially followed by a decrease, consistent with the trends shown in Figure 2.6 and Figure 2.7 Furthermore, both NEC and true event rate increase slower for the time resolution of 250 ps, compared to other values. This can be attributed to that fact that due to the small diameter of the system, the actual arrival time difference of 511 keV photons (less than 500 ps) may have exceeded the time window setting (e.g. 0.8x250 ps=200 ps). As a result, a portion of true events will have also been rejected and results in a reduced count rate.

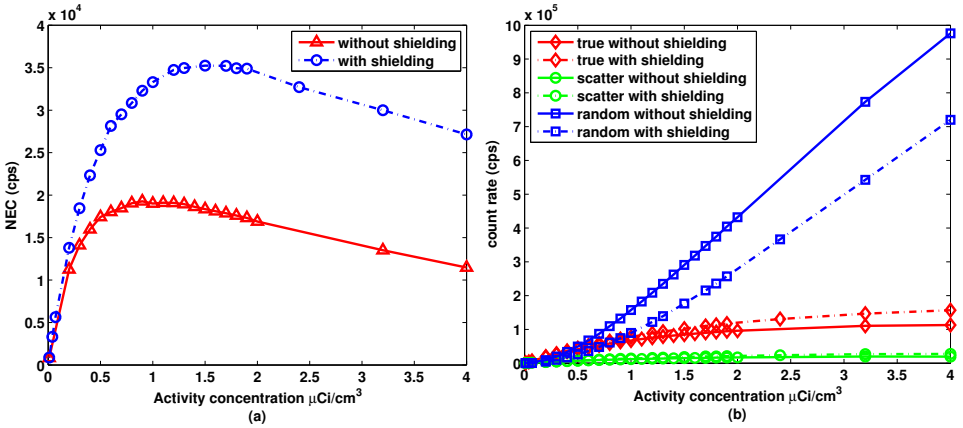

Figure 2.9: (a) NEC result as a function of activity concentration with and without lead shielding. (b) Coincidence count rate (including true, scatter, random events) as a function of activity concentration with and without lead shielding.

The dependency of NEC on activity concentration in breast is shown in Figure 2.9 (a). The activity concentrations in other organs were scaled accordingly as shown in Table 2.2. The time resolution was kept at 1 ns and the time window setting was kept at 1.2 ns based on results in Figure 2.8. It is found that the NEC reaches the peak value 19.2 kcps at 0.9 μCi/cm^3 but then decreases. Similar to the dependence of NEC on time window, this is mainly due to the large amount of random events from breast and adjacent organs outside of FoV as shown in Figure 2.9 (b). Moreover, the system dead-time also limits recording the coincidence events, when the activity is high, there will be more frequent occurrence of pulse pile-up where two events will be recorded in the same detectors within the light integration time, there will also be more than two coincidence events detected in the coincidence window(Daube-Witherspoon and Carson, 1991). Both of these cases will reduce the increase of system count rate.

The effect of shielding on the NEC performance for the proposed system is shown in Figure 2.9 (a). With the introduction of lead shielding (2 cm thick) around the detector rings, the NEC

peak reaches 35.3 kcps at the concentration of 1.5 μCi/cm^3. Under such condition of high activity, the shielding effectively reduces the random event by 39%, while increasing the true event by 21% as shown in Figure 2.9 (b), respectively. Nevertheless, for the current clinical protocol using Fludeoxyglucose (FDG) which is 3-5 kBq/mL (0.08-0.14 μCi/mL), the effect of lead shielding will not be apparent.

2.4.3 Spatial resolution

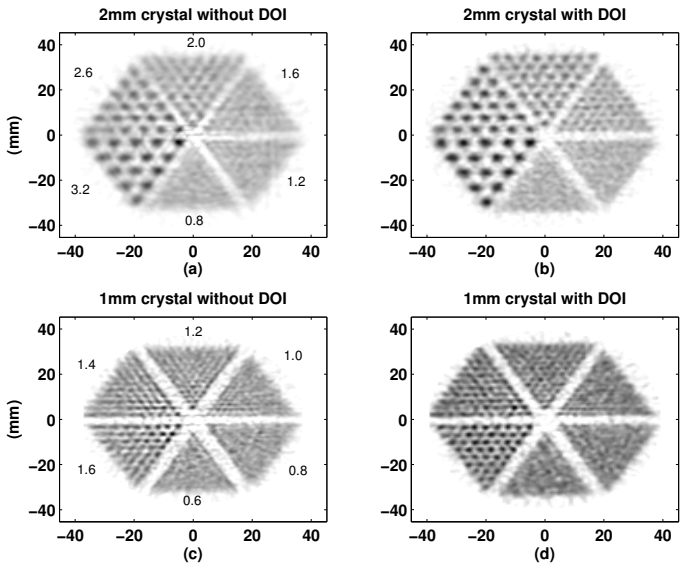

Figure 2.10: The in-plane images of hot Micro-Derenzo rod phantoms with cold background reconstructed by FBP with single-slice-rebinning. The pixel size is 0.5x0.5 mm^2.

Figure 2.10 shows the reconstructed Micro-Derenzo phantoms (the central slice) for the four different designs. It is observed that two designs (I and II) suffer noticeably from the parallax error effect towards the edge of FoV, due to the small diameter of the system configurations. While for the designs III and IV, 10 mm DOI capability can help mitigate the parallax effect. In addition, with the blurring effect of positron range excluded, 1 mm crystal design is able to clearly resolve the rods of 1.4 mm diameter and the 2 mm crystal design is able to clearly resolve the rods of 2.0

mm diameter. Due to limited spatial resolution, the image of smaller rods overlaps and appears blurred.

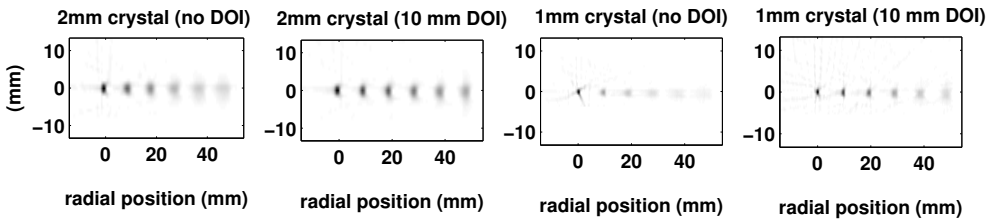

Figure 2.11: Quantitative study of DOI effect for different crystal designs. The in-plane image of six hot rods (diameter: 0.4 mm) with cold background extending from the center to the edge of the system are shown.

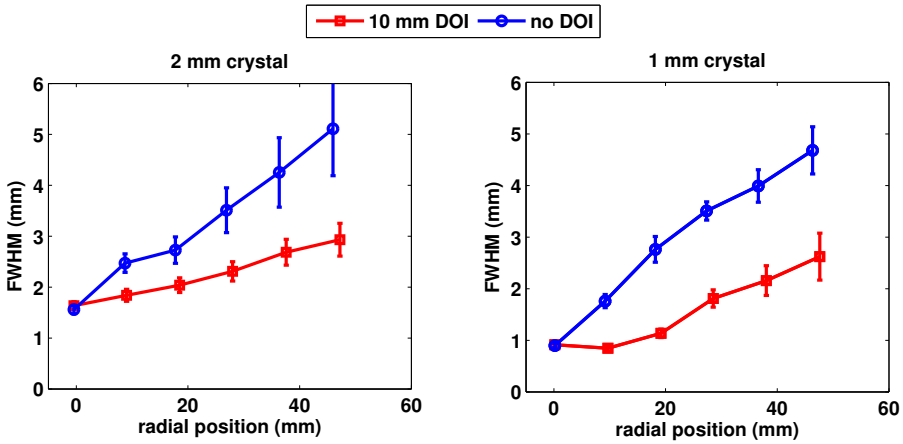

Figure 2.12: Reconstructed spatial resolution in FWHM as a function of radial position for different crystal designs obtained based on the images shown in Figure 2.11

The quantitative results of DOI study are shown in Figure 2.12, including rod sources with diameters equal to 0.4 mm separated by 1 cm distance. As described above, it is noticed that the spatial resolution degrades linearly along the radial offset, due to the parallax error. For the system I, the resolution is 1.56±0.08 mm FWHM at the center and 5.1±0.9 mm FWHM at 5 cm offset (degradation by a factor of 3.3). For the system II, the resolution is 0.89±0.04 mm FWHM at the center and 4.7±0.5 FWHM at 5cm offset (degradation by a factor of 5).

With the addition of 10 mm DOI, the FWHM improves especially towards the edge of the FOV. For example, for the system III, at the center of the FoV, the spatial resolution is 1.56±0.08, at the 5 cm offset from the center, the spatial resolution is 2.9±0.3 mm FWHM. In addition, though a slight improvement is observed between using 1 mm and 2 mm crystal elements with and without DOI when the radial offset is less than 3 cm, no considerable difference is observed when the radial offsets exceeds 4 cm between the two cases.

2.4.4 Contrast

All images in this section were reconstructed using 3D list-mode OSEM with 6 subsets. For the two cases having TOF capability (250 ps and 500 ps time resolution), an additional TOF kernel as shown in Equation. (3) and (4) was added. As shown in Figure 2.13 2.14, it is observed that the image quality improves with longer imaging time (7 mins vs. 2 mins), larger lesions (5mm vs. 3mm diameter) and higher activity concentration (10:1 vs. 5:1) all due to improved counting statistics. The same spheres in the images obtained for two TOF cases show improved detectability, compared to the non-TOF case.

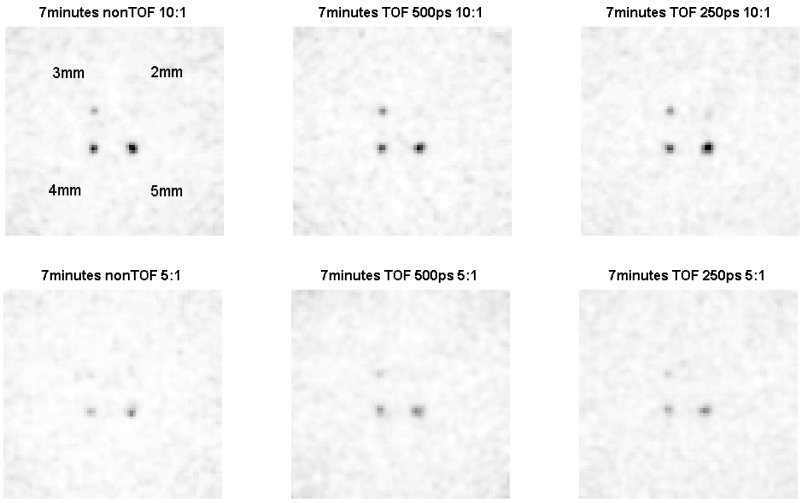

Figure 2.13: Reconstructed images of sphere lesions with different activity concentration (1st row: 10:1, 2nd row: 5:1) in the central slice, for non-TOF and TOF scenarios (500 ps and 250 ps time resolution). Scan time in the simulation is 7 minutes. The images were reconstructed using list-mode OSEM with 6 subsets and iteration numbers were optimized to achieve the highest CNR for 3 mm spheres (to be shown in Figure 2.15).

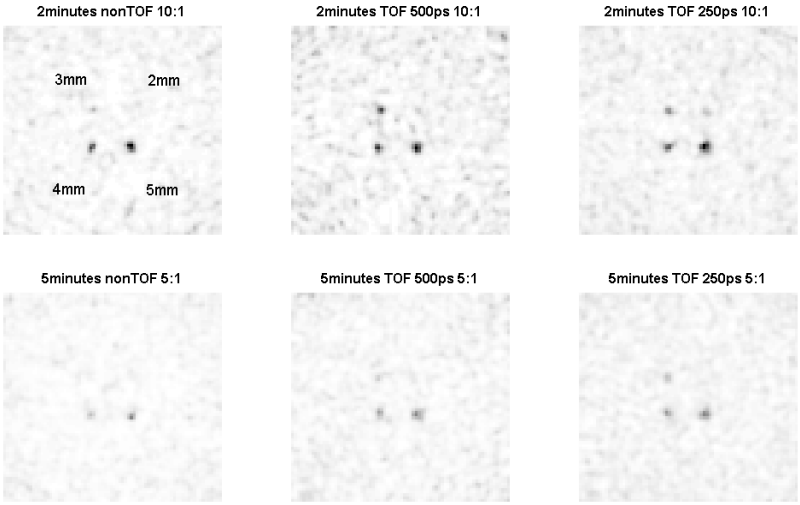

Figure 2.14: Illustration of image quality and lesion detectability for different activity concentration and scan time (2 minutes for 10:1 concentration, 5 minutes for 5:1 concentration). Other conditions are the same as those shown in Figure 2.13

Second, to optimize the image reconstruction for different spheres quantitatively, here, the dependency of CNR as a function of spheres size, TOF capability, iteration number and concentration ratio are shown in Figure 2.15. For example, when the imaging time is 7 minutes, the peak CNR values for the 5 mm sphere (10:1) are 42, 43.5 and 49 for three cases (non-TOF, 500 ps TOF and 250 ps TOF), respectively. In addition, it is found that the peak CNR value occurs at different iteration numbers for different sphere sizes. Simply applying a fixed iteration number would not achieve optimum CNR for spheres of different sizes. For the non-TOF case (activity concentration ratio 10:1), the iteration number corresponding to peak CNR is 3, 5 and 10, for the three sphere of 5 mm, 4 mm and 3 mm, respectively. Furthermore, for the 4 mm sphere (activity concentration ratio 10:1), the iteration number corresponding to peak NEC is 2 for non-TOF, 4 for the TOF (500 ps), and 5 for the TOF (250 ps). For all cases, increasing the iteration number beyond the optimum point results in the decrease of CNR due to the noise behavior of the OSEM reconstruction. No noticeable difference of such dependency on iteration number is observed between two activity

concentration ratios (10:1 and 5:1).

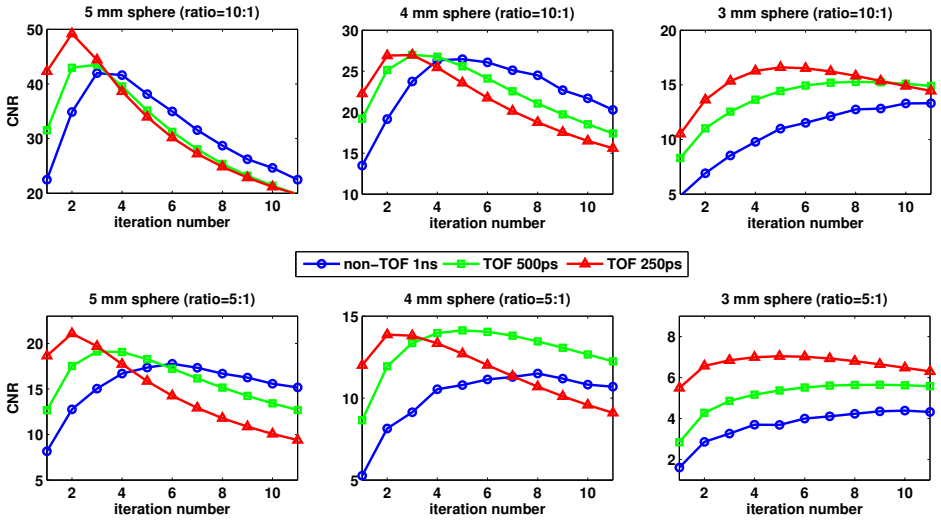

Figure 2.15: CNR results as a function of iteration number for lesions with different sizes (5 mm, 4 mm and 3 mm) and activity concentration ratios (10:1 and 5:1) for non-TOF and TOF scenarios.

Third, to answer the question what is the minimum required time to visualize a tumor of given size, the CNR value are studied as a function of scan time, as shown in Figure 2.16 The threshold value of 4 based on Rose criterion is also plotted as a reference to evaluate lesion detectability semi-quantitatively. The two TOF cases result in higher CNR than that of the non-TOF system at the same scan time. For instance, to detect the 3 mm sphere (activity concentration ratio 10:1), the non-TOF case requires more than 2 minutes to reach the threshold while both TOF cases can do in about 1 minute. To detect the 3 mm sphere (activity concentration ratio 5:1), the non-TOF case requires 7 minutes compared to 5 minute for the TOF-system cases. These results are also consistent with the qualitative analysis of images shown in previous sections. Referring to Figure 2.16, the 3 mm sphere can only be clearly visualized in two TOF cases in the following two conditions: 1) a scan time of 2 minutes and activity concentration ratio 10:1; 2) a scan time of 5 minutes and

activity concentration ratio 5:1.

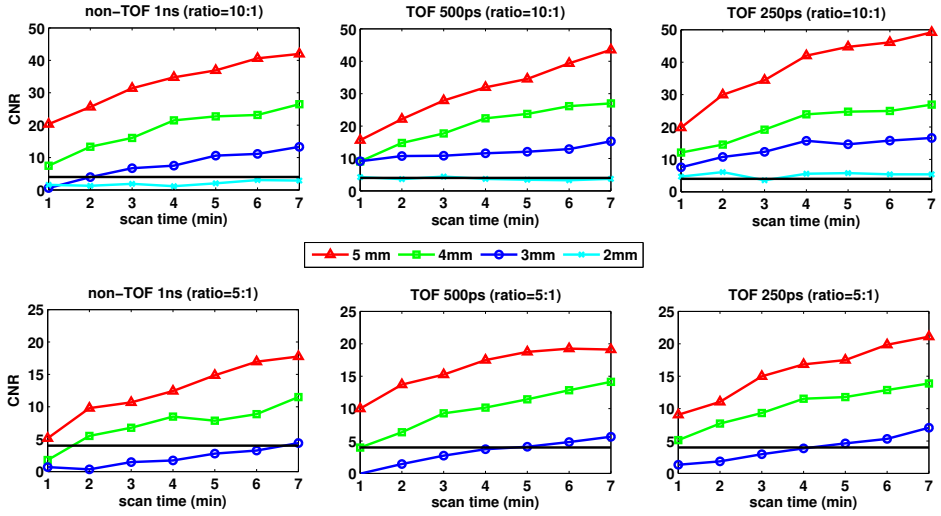

Figure 2.16: Study of CNR as a function of scan time for lesions with different sizes (5 mm, 4 mm and 3 mm) and activity concentration ratios (10:1 and 5:1) for non-TOF and TOF scenarios.

2.5 Discussion

In this work, our main focus is to evaluate the feasibility of designing a cost-effective breast dedicated PET insert, which could be integrated with MR breast coil. The system is expected to provide both structural and functional information with high sensitivity and specificity for breast cancer diagnosis, biopsy and surgery guidance. The simulation helps us optimize system design and answer several questions: 1) what is the minimum number of rings required to achieve sufficient counting statistics within reasonable imaging time? 2) What is the dependency of NEC performance on energy/time resolution, energy/time window settings, injected activity concentration and shielding design? 3) What is the effect of fine crystal element (1 or 2 mm) and DOI resolution (10 mm) on spatial resolution, resolution uniformity, and lesion detection? 4) How long will it take to achieve sufficient CNR for detecting small lesions of typical lesion-to-background activity ratio?

We found the relationship between photon sensitivity and the number of PET detector rings. The photon sensitivity shows dependency on spatial position across the FoV, due to both solid angle coverage and gaps in-between detector blocks. As shown in Figure 2.3, for a system of about 3.3 cm height (4 rings using 1x1x20 mm^3 crystals (8x8 array), or 3 rings of rings using 2x2x20 mm^3 crystals (5x5 array), the sensitivity is able to reach 4.7% and 6.2% respectively. This is comparable to other breast-dedicated PET systems utilizing novel geometry configurations ranging from 4.3% to 6.8% (Raylman *et al.*, 2008)(Abreu *et al.*, 2007)(Wang *et al.*, 2006).

This is a big improvement compared to the whole body PET systems, which have sensitivity about 1% (Habte *et al.*, 2007), To further improve the photon detection sensitivity, the number of the PET detector rings could be adjusted based on results in Figure 2.3, considered together with the requirements of imaging time and lesion contrast.

Another important finding is the effect of time resolution on NEC performance. First, it is found that by utilizing PET detector of improved time resolution down to 250 ps or 500 ps, we can improve NEC performance for the propose system, as shown in Figure 2.6. Second, for different time resolution values, the peak NEC occurs with different time window settings, different from typical settings in whole body imaging in which the time window is set to be twice the time resolution. For 250 ps time resolution, the peak NEC value is to be achieved at a time window of 500 ps. When the time window is set to be too small (250 ps) or too big (4 ns), it does not provide optimum NEC performance, as shown in Figure 2.6. Such dependency is mainly due to two reasons associated with small diameter geometry: 1) high random event from breast and adjacent organs; 2) minimum time required for 511 keV photons to travel through patients. Third, a lead shielding of 2 cm thickness is able to improve NEC, particularly for higher injected concentration (e.g., from 19.2 kcps to 32.3 kcps at 0.9 μCi/mL). Though such improvement is not considerable with the current clinical breast PET protocol which is 3-5 kBq/mL(MacDonald *et al.*, 2009) (0.08-0.14 μCi/mL), it may be beneficial for application using more specific radiopharmaceuticals such as FLT (Fluorothymidine) for which the injected activity can be increased to achieve reduced imaging time and improve image

quality in teams of SNR and CNR.

The resolution study shows that another benefit of the proposed system is high spatial resolution, capable of resolving spheres of 1.4 mm diameter (1 mm crystal element) and 2.0 mm diameter (2 mm crystal element) around the center of FOV. The spatial resolution of clinical whole-body PET systems are typically 5∼6 mm FWHM, limited by the large crystal dimension (4-5 mm). As a result, the proposed system is expected to be able to detect small lesions inside the breast, given enough counting statistics. The results of DOI study indicate that the incorporation of DOI resolution to the detector design is required to mitigate parallax error and improve resolution uniformity. For the crystal design IV, the two layer design (10 mm DOI) improves the resolution from 4.5±0.8 mm (design III) down to 2.9±0.5 mm at a radial offset of 5cm, compared to 0.88±0.01 mm at the center of FOV. We believe such resolution uniformity might be sufficient for clinical applications.

We studied the lesion detectability based on the crystal design IV. The images were compared between the non-TOF case and two TOF cases (500 ps and 250 ps time resolutions). The results indicate that the TOF capability could help lesion detection with less iteration number. For the 3 mm diameter lesion, the TOF capability is able to detect it within 1 minute for the activity concentration ratio of 10:1 (Figure 2.16) and within 5 minutes for the activity concentration ratio of 5:1 (Figure 2.16). In addition, to achieve best CNR, the iteration number needs to be optimized depending on the size of lesions to be detected (Figure 2.15). However, it remains a challenge for the proposed system (design IV, 2 mm crystal element) to detect lesions with diameters less than 2 mm within 7-10 minutes by extrapolating the results shown in (Figure 2.16), mainly limited by spatial resolution, photon detection sensitivity and NEC performance. Furthermore, it should be pointed out that due to the small size of the proposed system, the potential benefit of TOF is to be less noticeable as compared to a whole-body PET system. On the other hand, to achieve superior timing resolution less than 250-500 ps might beyond the current limit of currently available fast scintillators, photodetectors and readout electronics.

Regarding the selection between 1 mm and 2 mm crystal dimension in the system, several trade-offs need to be considered. A major question to be further answered is that based on high resolution anatomical information from MR images, the value of integrating PET is mainly to provide complementary functional/physiological information and improve specificity of cancer detection. As a result, this would make a high resolution PET system less a necessity. In the meantime, though 1 mm crystal design would yield superior spatial resolution, it poses several challenges on system design. First, in order to deploy 20 mm thick crystals to enhance photon detection sensitivity, the 1 mm crystal design suffers from increased dead area in-between crystal elements and poor packing fraction, compared to 2 mm crystal design. Second, the comparison with respect to spatial resolution between two designs shows that the improvement is quite limited within the region 4 cm diameter, at 4 cm both design III and IV have spatial resolution of 2.3 mm FWHM. Third, in the contrast study, if a pixel size is set to be even smaller (0.5 mm) according to the 1 mm crystal element compared to that (1 mm) of the 2 mm crystal design, more counting statistics and/or longer imaging time would be required to obtain the same SNR(Peng and Levin, 2010). Fourth, 1 mm design would result in degraded energy resolution and time resolution, due to reduced light collection. This will consequently degrade NEC performance. Furthermore, it seems not to be likely to develop an array of 1x1x20 mm^3 LYSO using SiPMs of 3 mm pixel size unless SiPMs of smaller pixel size around 1-2 mm are to be used.

To conclude, the simulation results in this work will help us optimize system design. Once the system is built, it is expected to have several impacts on breast cancer imaging. First, high resolution of the PET insert (2 mm) will allow earlier and more accurate detection of malignant lesions, with both high sensitivity (combined with MR images) and specificity. Though PET will not be used for routine screening due to radiation exposure, it may play an important role in screening for patients with high-risk of breast cancer and/or with radio-dense breast tissues. Second, it will help guide biopsy by improving the accuracy of localizing and sampling malignant tissue, as well as avoiding unnecessary and invasive procedures. Third, early detection of focal label accumulation

with PET will help improve the accuracy of identifying local recurrence post-treatment in conjunction with the use of functional MRI techniques, which is extremely important for patients after post-mastectomy. Finally, the system is expected to facilitate the monitoring of treatment response for chemotherapy/radiation therapy immediately after the first few cycles of treatments, by using complementary information from PET and MR.

Chapter 3

GPU based Image reconstruction with resolution modeling

The simulation results in chapter 2 indicate that for the proposed breast-dedicated PET/MR system, the spatial resolution suffers from parallax error and significant degradation towards the edge of the FOV. Such degradation can be mitigated with the incorporation of dual DOI layer design (10 mm), at the cost of increasing detector/system complexity. This chapter focuses on studying the feasibility of using an image reconstruction method (ordered-subset expectation maximization, OSEM) with resolution modeling taken into account to improve the spatial resolution for the proposed system, requiring no physical DOI design (i.e. dual-layer readout design) in place. A majority portion of this chapter has been implemented with graphic processing unit (GPU). More details about OSEM PET reconstruction and GPU platform are provided in Appendix for reference.

3.1 Resolution modeling in OSEM

In a PET system, a measured coincidence event detected by a detector pair provides the information about the distribution of the radionuclide in the body (i.e. line-of-response, LOR). In terms of

mathematical expression, it could be written as(Alessio and Kinahan, 2006):

$$N = Px \tag{3.1}$$

where N is the number of the coincidence events (also referred to projection), x is the tracer distribution discretized in the image space. P is the system matrix associated with the probability that a pixel (x_i) contributes to a given projection (N_j). The image reconstruction process is an inverse problem and by accurately modeling the system matrix, we are able to take into account of several resolution blurring factors such as positron range and detector response, which is not possible for filter-back projection (FBP).

The system matrix is determined by a number of different physical factors, which could be expressed as a combination of several probability matrix terms (Qi *et al.*, 1998), which follows the expression:

$$P = P_{det.sens} P_{attn} P_{positron} P_{geom} P_{det.blur} \tag{3.2}$$

Here the term $P_{det.sens}$ represents the detector normalization matrix, which could be measured by a uniform cylinder source. P_{attn} is the attenuation matrix which is used for attenuation correction. $P_{positron}$ consider the effect of the positron range. P_{geom} is the geometrical matrix which represents the geometrical probability that a photon passing through a voxel is detected by a detector pair, where the position of the detector pairs is only chosen to be the front surface or the center of the detectors. $P_{det.blur}$ consider the detection of photon in other positions in the detector other than the front surface or the center, which could be caused by photon non-collinearity, crystal scatter, crystal penetration and sinogram blurring.

The dominant factor of the system matrix is the geometrical matrix Aguiar *et al.* (2010), which is usually calculated by the ray tracing method such as the Siddon's method Siddon (1985). It

is widely deployed to the iterative reconstruction process due to its simplicity, however it is not able to take into account other factors which are very important for a high resolution PET system as described above. For the proposed system, a method that is capable of modeling the accurate system matrix and parallax error (term $P_{det.blur}$)is desired.

There are 3 different approaches to obtain the system matrix (P) for resolution modeling in PET OSEM reconstruction: 1) experimental measurement; 2) Monte-Carlo simulation; and 3) analytical modeling. The strengths and limitations of these methods are briefly discussed below.

Experimental measurement (Alessio *et al.*, 2010) (Wiant *et al.*, 2010) (Panin *et al.*, 2006)is conducted by placing a point source inside the FOV and measure the total counts to be detected for each detector pair (i.e. all possible LoRs). Once sufficient amount of counts are acquired (10^7 coincidence events), the point source is then placed in another position to obtain point spread function. Such method account for all the factors except attenuation that contribute to the system matrix, however the measurement requires time consuming acquisition at multiple locations. This task is in particular a challenge for resolution modeling for a PET system with large image space. For a standard whole body PET scanner, the data acquisition for a single 2D transaxial slice takes roughly 7 hours and requires 1.3 Gb for data storage (Wiant *et al.*, 2010).

Instead of measuring with a point source, the second method derives the system matrix using Monte Carlo simulation (Rafecas *et al.*, 2004)(Alessio *et al.*, 2006). Compared to direct measurement, the simulation setup is relatively easier, and it also offers flexibility to optimize designs for novel PET systems (i.e. geometry). However, such method has a very high demand on the computation resources and can also suffer from statistical noise. In addition, compared to the direct measurement, this method may face limitations in taking into account all the factors described in formula 3.2. For example, existing simulation softwares such as GATE lacks the capability to accurately model the crystal identification process inside an Anger-logic based PET block detector(Panin *et al.*, 2006).

Finally, the system matrix could also be modeled by analytical calculation(Liang, 1994)(Strul *et al.*, 2003), where the probability of detection of a pair of photons in one direction is modeled and integrated to all possible directions. Without modeling all factors, this method assumes that the system matrix is mainly dominated by one factor over others. For example, for the small animal PET scanners where the parallax error is dominated, other factors such as intercrystal scatter, positron range and photon non-collinearity can be ignored. The advantage of this method is that it does not suffer from statistical noise and requires no time consuming measurement. This chapter will focus on the analytical calculation of system matrix.

3.2 Coincidence detector response function (CDRF)

In this work, we are focusing on the analytical approach that only models the parallax error (due to crystal penetration effect) for system matrix calculation. Two reasons for such simplification are provided below. First, the main motivation is to investigate if the use of resolution modeling is able to help the proposed system to mitigate parallax error so that there is no need to introduce physical DOI. Second, the proposed system is configured in the 3D acquisition mode and the dimension of the system matrix is large, which requires the calculation and storage of the system matrix to be computed on the fly inside the GPU. As a result, the complexity of the modeling needs to be reduced.

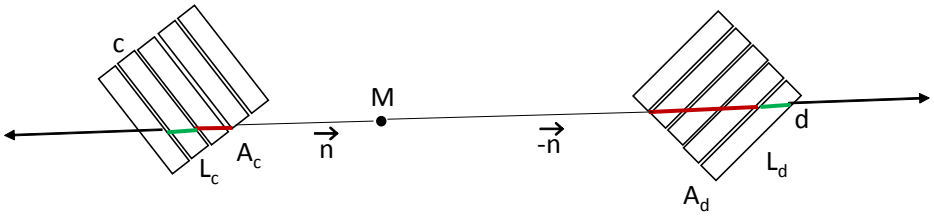

Figure 3.1: The 511 KeV photons emitted at point M along direction $-\vec{n}$ is attenuated with A_d in length and detected with L_d in crystal d.

The crystal penetration effect is shown in Figure 3.1, where before the photons are detected in a crystal d, they may penetrate to the adjacent crystals with a distance of A_d. The probability that

the photon is detected in the crystal with penetration is less than the crystal without penetration. To consider this effect, mathematical model that calculate the spatial variant probability distribution of detection is proposed which is called detector response function, defined by the following equation(Lecomte *et al.*, 1984)(Schmitt *et al.*, 1988)(Liang, 1994)(Strul *et al.*, 2003).

$$\varepsilon_{cd}(M) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} p_c(M,\vec{n})p_d(M,-\vec{n})sin\phi d\phi d\theta \tag{3.3a}$$

$$p_c(M,\vec{n}) = \{1 - exp[-\mu_c L_c(M,\vec{n})]\}exp[-\mu_c A_c(M,\vec{n})] \tag{3.3b}$$

where $\varepsilon_{cd}(M)$ is called the coincidence detector response function (CDRF) which is the detection probability for both crystal c and crystal d with the annihilation point located at M. $p_c(M,\vec{n})$ is the detection probability for photons originated from M along path $\vec{n}$ in the crystal c considering attenuation before entering crystal c. The interaction length and the attenuation length are L_c and A_c respectively. For a photon pair that originated at point M, all the possible path that is detected by crystal pair c and d is considered and the probability of detection for each path is integrated to represent the probability that the photons are detected in crystal pair c and d.

3.3 GPU accelerated resolution modeling

A number of published works have implemented the coincidence response function calculation to model the crystal penetration effect (Strul *et al.*, 2003)(Rahmim *et al.*, 2008)(Yamaya *et al.*, 2005). Most of the works pre-calculate and save the system matrix to be used in the image reconstruction process. As the system matrix connect the image space and projection space, it requires large amount of memory to be stored even though it could be compressed by using the geometry symmetry condition(Ortuño *et al.*, 2010)(Zhang *et al.*, 2010). To reduce the memory storage, the system matrix could be calculated on the fly during the reconstruction. However, it is very computational intensive that is not practical to be implemented to be calculated in serial with CPU. (Pratx and Levin, 2011) proposed a method to calculate the system matrix with CDRF modeling on the GPU device using OpenGL and cg shading libraries. The reconstruction time for processing 1 million

LORs is reported to be only 29.9 second after performing the additional calculation of term P (equation 3.2) which only increased the computational time by a factor of 10. However, after modeling the CDRF, the spatial resolution is found to be more uniform where the variation in reconstructed sphere size reduced from 0.5 mm to 0.07 mm RMS for 1.75 mm spheres. The CDRF also improves the contrast recovery performance. The contrast recovery for 8mm sphere is increased from 85.9% to 95.8%.

In this project, we propose to develop a GPU-based analytical method to accelerate the CDRF calculation and image reconstruction with resolution modeling. The proposed method mainly differs from the previous works in the following aspects. First, the proposed method considers the crystal penetration effect among different slices in 3D domain for any arbitrary system geometry, which was not available in the previous works(Pratx and Levin, 2011)(Rahmim *et al.*, 2008).

Second, a multiple-ray tracing method is applied to approximate the solid angle coverage of each LOR instead of calculating the solid angle boundaries which is a challenging task to be performed on the basis of individual voxels in the image space. For example, fewer voxels located in the center FOV are intersected by the LOR compared to those located at the edge. Third, the proposed model minimizes any difference of work load among different LORs, which is a major benefit with respect to efficient implementation with GPU. For instance, the work load for the LORs having larger solid angle coverage is larger than those have smaller solid angle coverage. Fourth, our method is performed on the GPU using the Nvidia CUDA toolkit which is easier to implement and maintain than OpenGL and cg shading libraries, a major trend found in the recent research on GPU-based image reconstruction(Eklund *et al.*, 2013)

Our proposed method is based on the concept of virtual DOI layers. In other words, instead of modeling a coincidence detector pair as a single LOR, we model it as multiple LoRs resulted from virtual DOI layers (note that no physical layer is present). With the addition of virtual DOI layers, the CDRF can be modeled with improved accuracy, when compared to the scenarios of both

no DOI case and DOI modeling using a spatial invariant kernel as shown in chapter 2. Our main hypothesis there is that if we can virtually create the virtual DOI layers inside a long crystal and calculate their respective contribution to the overall CDRF, the achievable spatial resolution after the resolution modeling is expected to be better than that of the case without DOI. The addition of virtual DOI layers split a single LOR to several sub-LORs (i.e. n=4) with the number equal to the square of the number of virtual DOI layers (i.e. N=n^2=16). Consequently, each of the sub-LORs is processed simultaneously with GPU with a weighted probability distribution for each sub-LOR taken into account (see Figure 3.4). The probability of each of the sub-LOR is analytically calculated (see section 3.4).

Several assumptions were made to simplify GPU implementation in our work. Firstly, the CDRF calculation method described in section 3.2 depends on source location (i.e. solid angle coverage is position dependent), however, in the proposed method, we assume that the probability of detection for a given source along the multiple sub-LOR is identical. Second, the calculation is further simplified when the probability of detections for a sub-LORs could be expressed as:

$$p_{c_i} = (1 - exp(-\mu L_{int}))exp(-\mu L_{atten}) \tag{3.4}$$

where L_{atten} is the crystal attenuation length from the crystal block surface to the center of the virtual DOI layer, and L_{int} is the crystal intersection length inside the DOI layer. For the LYSO crystal μ is $0.0869mm^{-1}$ (Melcher, 2000). This approximation is more robust when the crystal is divided into more subsections as the intersection length at any other positions is roughly equal to that passes through the center. Finally, sub crystal layers are only created in the crystal depth direction because the width of the crystal surface (2 mm) is much smaller than the thickness of the crystal (20 mm). The final system matrix is then calculated as the sum of the system matrices of each sub-LOR determined by the product of the geometrical matrix defined by a shift invariant Gaussian kernel and the probability of detection described in equation 3.4.

3.3.1 Crystal intersection length calculation

This section presents the method of calculating the crystal intersection length. Considering a photon passing through the center of a virtual DOI layer which in this work is a rectangular box as shown in Figure 3.2.

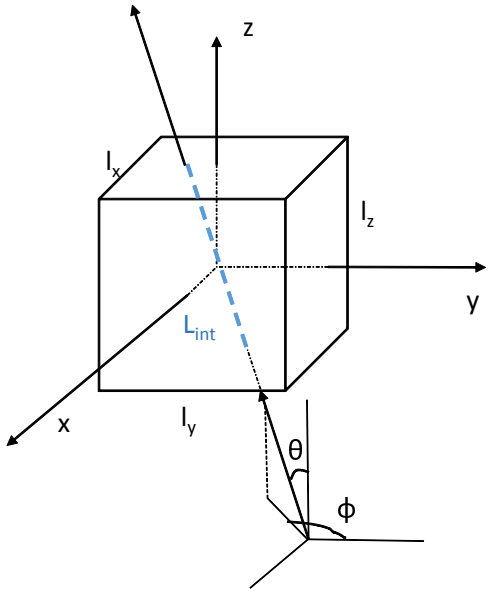

Figure 3.2: Crystal intersection length calculation

As the line passes the center, the total intersection length is twice the length from the surface intersection point to the box center. Now the problem of the calculation is how to select the surface. Assuming the size of the rectangular box is $L_x \times L_y \times L_z\ mm^3$, a coordinate system could then be set with the origin point set to be the center of the box, the six surfaces could be expressed as $(\pm 0.5L_x, y, z)$, $(x, \pm 0.5L_y, z)$, $(x, y, \pm 0.5L_z)$, where $(-0.5L_x < x < 0.5L_x), (-0.5L_y < y < 0.5L_y), (-0.5L_z < z < 0.5L_z)$. Reversely thinking, the intersection length could be determined by tracing a line from the center to the box surface. In terms of spherical coordinate, the intersection

length r has the relationship with the angle θ, ϕ and the surface could be expressed as:

$$rsin\theta cos\phi \leqq \frac{L_x}{2} \tag{3.5a}$$

$$rsin\theta sin\phi \leqq \frac{L_y}{2} \tag{3.5b}$$

$$rcos\theta \leqq \frac{L_z}{2} \tag{3.5c}$$

Any r value that meet the above condition is inside the box, the maximum allowable r value is the intersection length from the box center to any of the surface given the angle θ and ϕ. Therefore the intersection length inside the rectangular box is:

$$L_{int} = min(\frac{L_x}{sin\theta\cos\phi}, \frac{L_y}{sin\theta sin\phi}, \frac{L_z}{cos\theta}) \tag{3.6}$$

3.3.2 Crystal attenuation length calculation

Before entering the crystal being detected, the 511 keV photon could penetrate and interact with the adjacent crystals, the attenuation length should also be considered. To calculate the crystal attenuation length, other than the crystal, the geometry of the crystal blocks should also be taken into account.

Similar to the calculation of crystal intersection length, a coordinate system could be established with the origin set to be the center of the virtual DOI layer. However, in this case, the boundary of the crystal blocks is chosen instead of the virtual DOI layers. The symmetry condition where the total length could be calculated by doubling the length between the center to the surface could no longer be used.

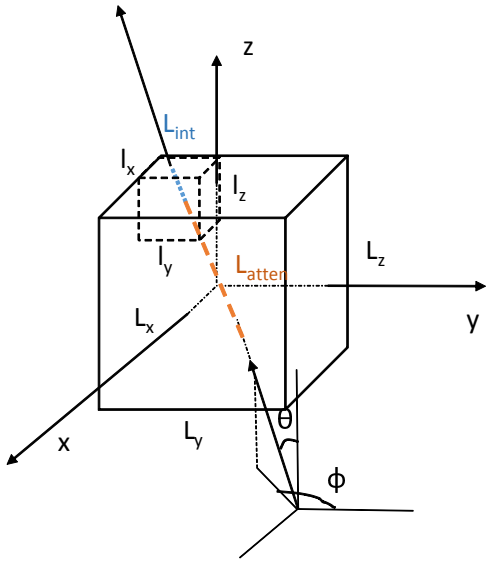

Figure 3.3: Crystal attenuation length calculation

As shown in Figure 3.3, The length, width and height of the crystal blocks are L_x, L_y, and L_z respectively. There are six surfaces of the detector blocks. Therefore the selection of front or back of the surface to calculate the crystal attenuation length should be determined first. The direction of the incoming photon and the surface could be expressed using vectors, the front or back surface could then be determined by the following condition:

$$(x_1 - x_0, y_1 - y_0, z_1 - z_0) \cdot (\ (-0.5L_x, 0, 0)\ R_{module}) > 0\ (front\ yz\ surface) \tag{3.7a}$$

$$(x_1 - x_0, y_1 - y_0, z_1 - z_0) \cdot (\ (0, -0.5L_y, 0)\ R_{module}) > 0\ (front\ xz\ surface) \tag{3.7b}$$

$$(x_1 - x_0, y_1 - y_0, z_1 - z_0) \cdot (\ (0, 0, -0.5L_z)\ R_{module}) > 0\ (front\ xy\ surface) \tag{3.7c}$$

where the photons are coming from (x_1, y_1, z_1) to (x_0, y_0, z_0), and R_{module} is is the rotation matrix of the crystal module. Then the attenuation length could be expressed as:

$$L_{atten} = min(\frac{L_{yz}}{sin\theta \cos\phi}, \frac{L_{xz}}{sin\theta sin\phi}, \frac{L_{xy}}{cos\theta}) - L_{int}/2 \tag{3.8}$$

where L_{yz}, L_{xz}, l_{xy} in the back surface are $0.5l_x$, $0.5l_y$, $0.5l_z$ and in the front surface are $L_x - 0.5l_x$, $L_y - 0.5l_y$, $L_z - 0.5l_z$ respectively.

3.3.3 Model Validation

A. Method

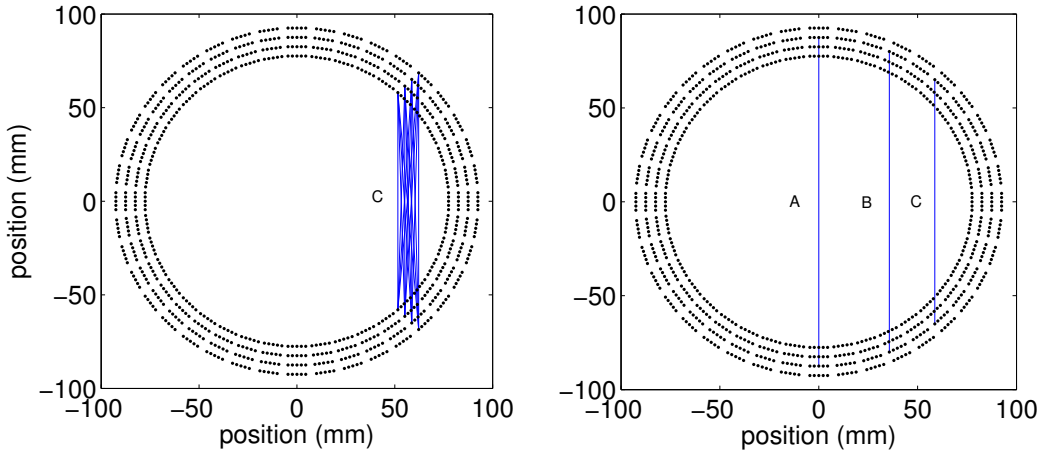

Figure 3.4: Three LORs (A,B,C) are chosen for validation analysis shown on the right. The LOR C is split to 16 sub-LORs shown on the left for illustration of the effect of multiray projection.

The essential concept of the proposed method is shown in Figure 3.4. The system has a diameter of 150 mm and comprise 40 blocks of detector per ring, where the single LOR C from two crystals at the edge of FOV is splited into 16 sub-LORs originating from 4 virtual DOI layers on either side. In our study, different numbers (n) of virtual DOI layers were also tested (n=2,3,4, etc).

To validate the proposed method, the comparison was made between our analytically calculated system matrix and the system matrix obtained from Monte Carlo simulation. The simulation was performed using the GATE simulation toolkit(Jan *et al.*, 2004), where a single point source of 0.1mm diameter was placed at multiple locations along the X axis (see figure 3.4 for coordinate) ranging from -7 mm to 7 mm for LOR A, 26 mm to 42 mm for LOR B, and 48 mm to 62 mm

for LOR C with a step size of 0.25 mm. Two other y positions (20 mm and 50 mm) were also tested. The normalized distribution of total coincidence counts for a selected pair of detectors at different source positions were recorded to provide the system matrix for the corresponding x and y position values. It should be pointed out that in our current study, the source locations of different z values (perpendicular to the XY plane) were not considered due to the fact that our proposed ring insert has a very small axial FOV (33.6 mm) compared to the transaxial FOV which is 150 mm. Therefore, the effect of crystal penetration is not expected to be obvious in different slices.

B. Result

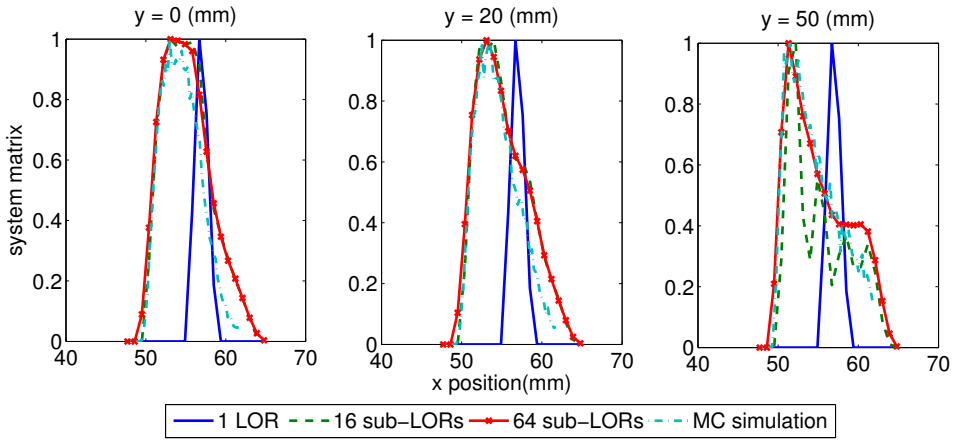

Figure 3.5: System matrix of LOR C plotted as a function of x position with y position set to be 0 mm (left), 20 mm (middle) and 50 mm(right). The result obtained from Monte Carlo simulation (cyan) are compared to the analytical calculation with 1(blue), 16(green), and 64(red) sub-LORs modeled.

Figure 3.5 shows the result of system matrix for LOR C using either Monte Carlo simulation or the analytical method described in section 3.3 with 1, 4 and 8 virtual DOI layers. There are 3 significant findings to be mentioned here. First, a position offset of approximately 5 mm (i.e. the peak location) is observed. This means that without using the multiple ray tracing method as we proposed, a single LOR with the spatially invariant Gaussian kernel does not reflect the true

distribution of detection probability along the profile of system matrix. Second, as the number of sub-LOR modeled increases from 16 (n=4) to 64 (n=8), the profile of system matrix obtained with the proposed method is approaching the profile obtained from Monte Carlo simulation (i.e. overlapping of the two profiles).

Finally, for locations of larger x values (towards the edge of FOV), the discrepancy between the results from the proposed method and the Monte Carlo method exhibits a weak dependency on y values (largest difference seen for y= 50 mm). A possible explanation for this is that the proposed method only considers the attenuation within a single crystal block ((i.e. block unit: 5x5 LYSO crystal) and has not yet taken into account of the penetration through adjacent crystal blocks. For those points towards the edge of FOV, the probability of such interaction across adjacent blocks is to be higher for larger y values compared to small y values. Such explanation is also supported by the results of system matrix profiles for LOR B when x ranges between 26 mm and 42 mm (Figure. 3.6), where the calculated system matrix better matches that from Monte Carlo simulation for all three y values since for this range of source locations, no significant interaction across adjacent blocks is expected.

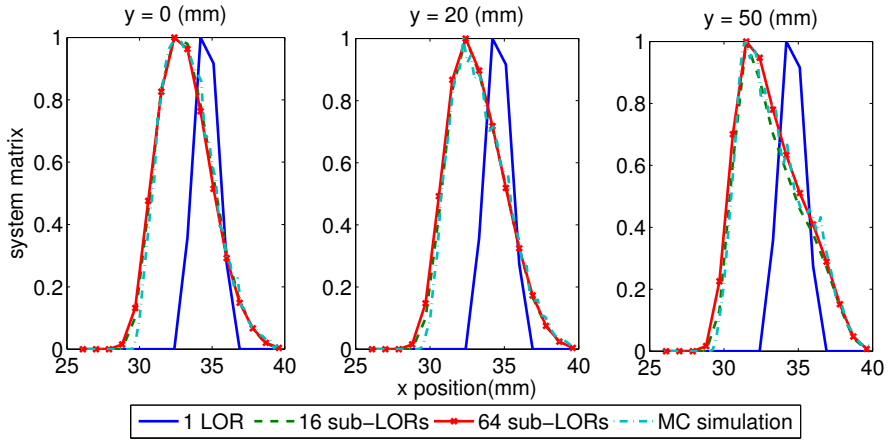

Figure 3.6: System matrix of LOR B plotted as a function of x position with y position set to be 0mm (left), 20mm (middle) and 50mm(right). The result obtained from Monte Carlo simulation (cyan) are compared to the analytical calculation with 1(blue), 16(green), and 64(red) sub-LORs modeled.

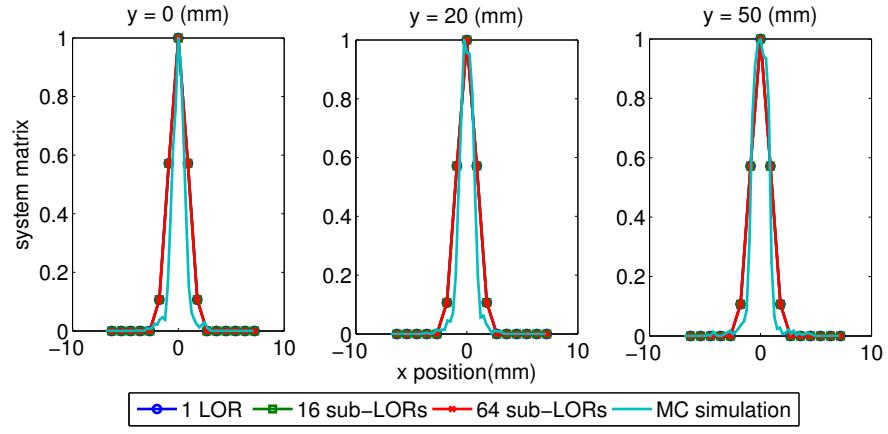

Figure 3.7: System matrix of LOR A plotted as a function of x position with y position set to be 0mm (left), 20mm (middle) and 50mm(right). The result obtained from Monte Carlo simulation (cyan) are compared to the analytical calculation with 1(blue), 16(green), and 64(red) sub-LORs modeled.

Another comparison is made among the three different methods for the LOR A with x ranging from -7 mm to 7 mm. In this case, no interaction through adjacent blocks is possible. All cases have a triangle-like distribution as consistent with previous work(Strul *et al.*, 2003) and no significant difference is found from the result. One feature to be mentioned is that the calculated system matrix using the multiray method has a slightly broader width compared to that obtained by Monte Carlo simulation. This is a consequence of using a shift invariant Gaussian kernel with a fixed FWHM set equal to the crystal width and creating sub-LORs only in the crystal depth direction resulting in no solid angle improvement as all of the sub-LORs overlap together.

The validation results in this section help us validate the accuracy of the proposed method. In the following section, we will further compare the differences of image quality quantitatively among these method, in terms of spatial resolution and contrast recovery.

3.3.4 Spatial resolution

A. Method

To test the potential improvement of spatial resolution using the proposed method, a phantom comprising 13 long rods with diameters of 0.2 mm and activity concentration of 1 Ci/cm^3 was used. The phantom was then placed in the central slice and the interval between two adjacent rods was set to be 10 mm. The resolution study includes the comparison among three scenarios, including: 1) a system without DOI layer modeled using a spatially invariant kernel (case 1); 2) a system with two built-in DOI layers modeled using a spatially invariant kernel (case 2); 3) a system without built-in DOI layers but with the proposed method applied for different number of virtual DOI layers (n=2,3,4) case 3. All the images were reconstructed using 6 subsets and 10 iterations.

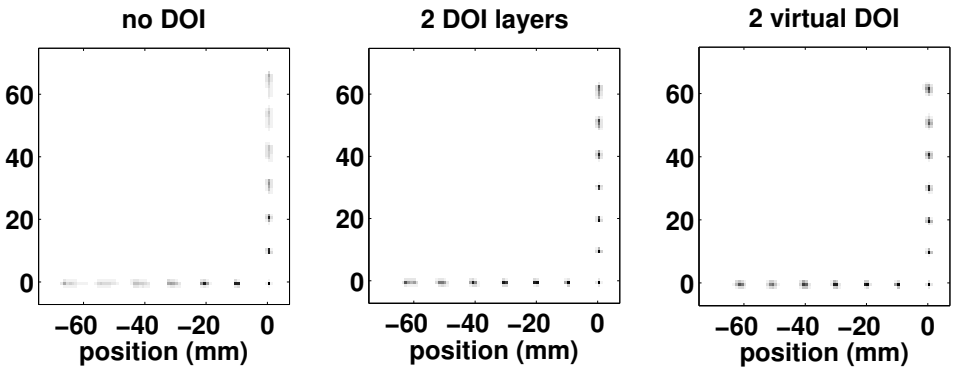

Figure 3.8: Image reconstructed for case 1 (left), case 2 (middle) and case 3 (right). All images are reconstructed with 6 subset and 10 iterations (voxel size $0.9 \times 0.9 \times 0.9mm^3$).

Figure 3.8 shows an example of the reconstructed images of the phantom. To quantitatively analyze the results and spatial resolution in terms of FWHM and offset, line profiles were drawn along X axis shown in Figure 3.8 and were fitted with Gaussian functions to derive the FWHM values of individual peaks. For case 1, the profile of the two rods on the very edge (X=60 mm and 50 mm) could not be fitted using Gaussian functions, with a R^2 equal to 0.65 from the fitting. Therefore, the FWHM analysis was also performed using a direct linear interpolation to validate the results from Gaussian fitting.

B. Result

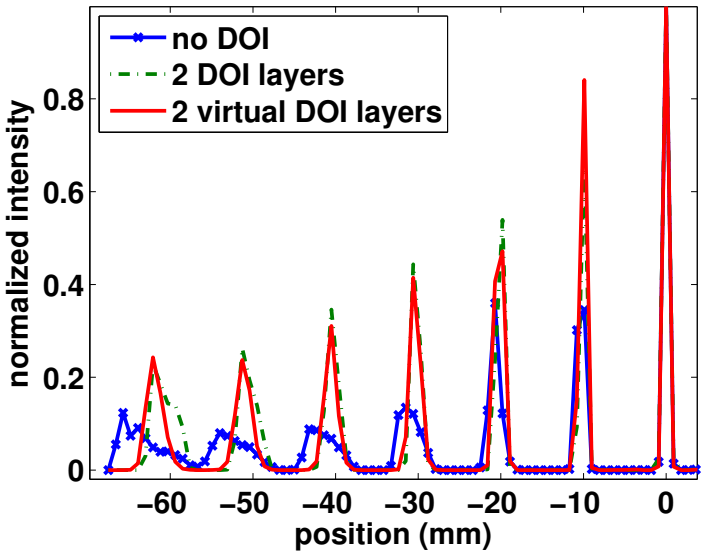

Figure 3.9: Line profile created along x axis of the images in Figure 3.8. The line profiles are fitted with Gaussian function to study the spatial resolution and peak position offset quantitatively. Note that the last two peaks (-50 mm, -60 mm) of the blue curve are excluded due to its non-Gaussian shape.

The profiles of the rod phantom are shown in Figure 3.9. As predicted, the spatial resolution degrades from the center of FOV toward the edge due to parallax error for all three cases. The quantitative measurements of FWHM were shown in Figure 3.10 (Gaussian fitting method) and Figure 3.11 (linear interpolation method). That's why two data points (60 mm and 50 mm) for the case 1 and the corresponding values are not shown in Figure 3.10. For other x values, good agreement is found between the two methods and only the results of X=30 mm are shown for illustration.

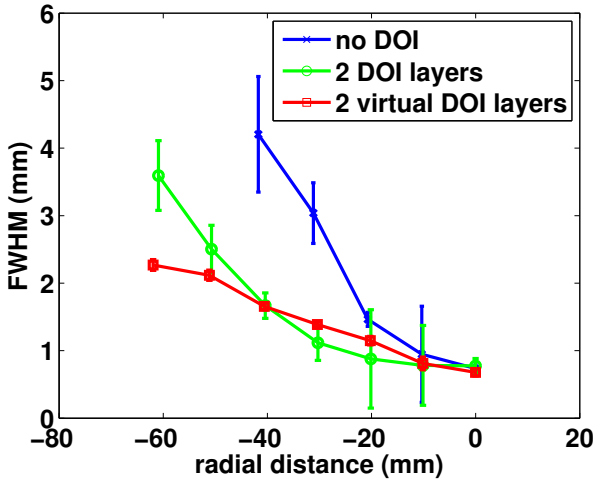

Figure 3.10: Spatial resolution are quantified as the FWHM of the fitted Gaussian functions for the line profiles in Figure 3.9. The last two peaks of the system without DOI and without resolution modeling (blue curve) is analyzed separately shown in Figure 3.11.

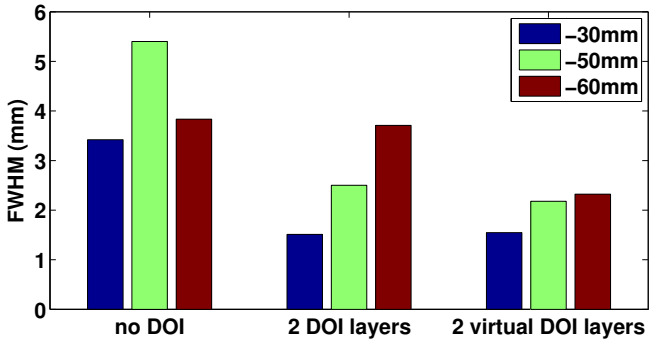

Figure 3.11: FWHM of the last two peaks (-60 mm and -50 mm) for all the profiles shown in Figure 3.9 are also calculated directly using linear interpolation instead of Gaussian fitting. The FWHM of the peak at -30 mm is also shown for comparison.

First, it is observed that case 3 (n=2) demonstrates improved spatial solution over case 1 and case 2, as shown in Figure 3.11 For the rod at 60 mm, the FWHM is 3.83 mm (case 1), 3.71 mm (case

2), and 2.32 mm (case 3), respectively. Due to the nature of linear interpolation, no uncertainly is included here. For the rod at 30 mm, the FWHM is 3.42 mm (case 1), 1.51 mm (case 2), and 1.55 mm (case 3), respectively. These results are consistent with those shown in Figure 3.10, obtained from Gaussian fitting.

Second, Figure 3.10 shows that for a given position, the worst spatial resolution is found for case 1. In addition, the spatial resolution is found to be more uniform for case 3, when compared to case 2 and case 1. For instance, the difference of the spatial resolution at radial distance of 20 and 40 mm for case 3 is 0.51 ± 0.08 mm which is smaller than the other two systems, where the difference for case 2 is 0.79 ± 0.75 mm and the difference for case 1 is 2.74 ± 0.86 mm.

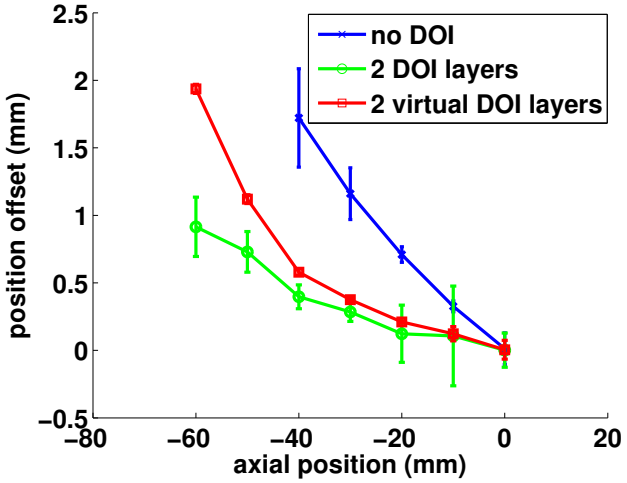

Figure 3.12: Position offset calculated as the absolute difference between the position of the fitted Gaussian peaks and the exact peak position. The last two peaks of the case 1 (blue curve) is analyzed separately shown in Figure 3.13

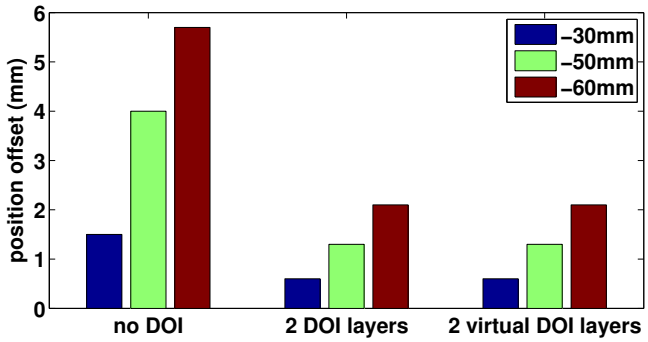

Figure 3.13: Position offset of the last two peaks (-60 mm and -50 mm) for all the profiles shown in Figure 3.9 calculated directly using linear interpolation instead of Gaussian fitting. The position offset of the peak at -30 mm is also shown for comparison.

Besides resolution FWHM, another parameter studied is the position offset, defined as the difference between the peak position from fitting procedure and the true position, which is also of critical importance in lesion localization along with FWHM. As shown in Figure 3.12 both case 2 and case 3 greatly reduce the effect of mis-positioning. The mis-positioning increases linearly as a function of radial distance and the most significant one is observed for case 1, up to 5.7 mm at X=60 mm. For case 2 and case 3, the maximum position shift is reduced to 2.1 mm.

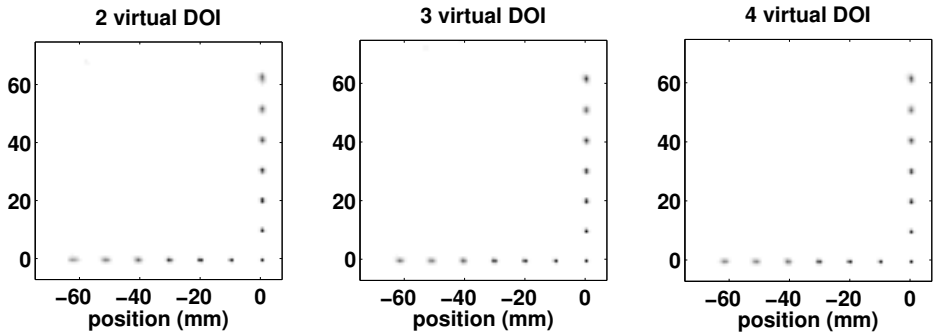

Figure 3.14: Image reconstructed with resolution modeling using 2 (left), 3 (middle), and 4 (right) virtual DOI layers. All images are reconstructed with 6 subset and 10 iterations (voxel size $0.9 \times 0.9 \times 0.9mm^3$).

Finally, comparisons were made to investigate if different number of virtual DOI layers (n=2,3,4) would impact the image quality, similar to the studies above. Figure 3.14 shows the reconstructed image of the phantom in the central slice. No significant difference is observed in the reconstructed images.

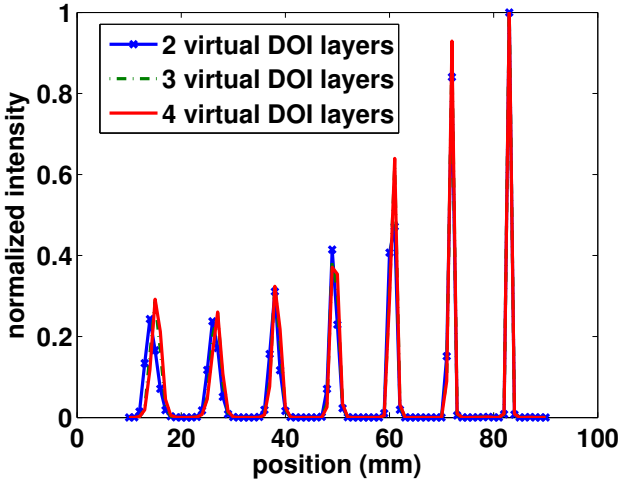

Figure 3.15: Line profile created along x axis of the images in Figure. 3.14. The line profiles are fitted with Gaussian function to study the spatial resolution and peak position offset quantitatively.

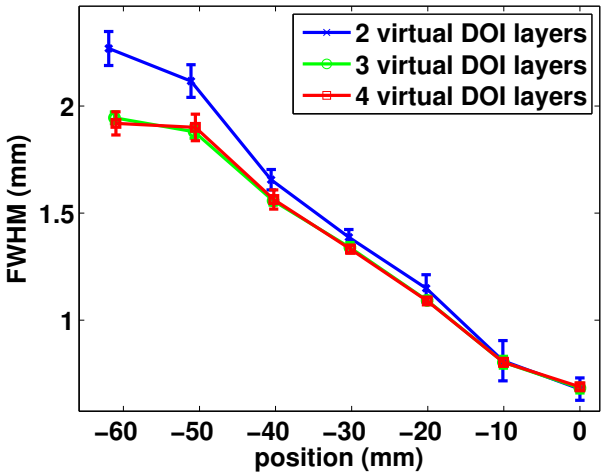

Figure 3.16: Spatial resolution are quantified as the FWHM of the fitted Gaussian functions for the line profiles in Figure 3.15

The results shown in figures 3.15 and 3.16 indicate that by increasing the number of virtual DOI layers from 2 to 3, the spatial resolution FWHM can be improved slightly. However, further increasing the number of layers has no significant impact, as seen from the results for n=3 and 4. For example, even for a position of 60 mm away from the center, the FWHM is 2.27 ± 0.08 mm (n=2), 1.94 ± 0.03 mm (n=3) and 1.92 ± 0.05 mm (n=4), respectively.

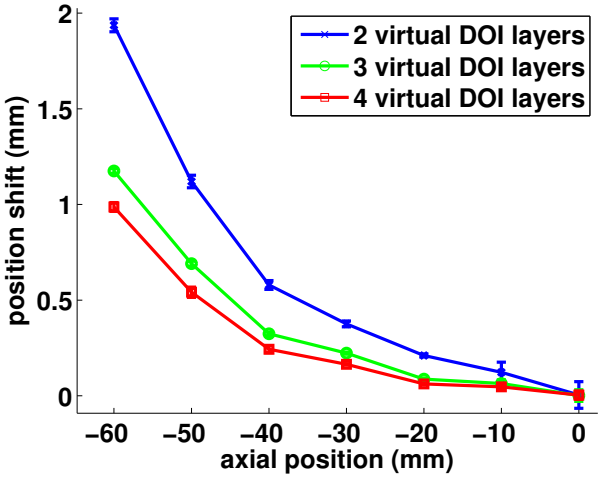

Figure 3.17: Position offset are calculated as the absolute difference between exact peak position and the the position of the fitted Gaussian peaks for the line profiles in Figure 3.15

Regarding the position offset shown in Figure 3.17, with more virtual DOI layers modeled, the less position offset was found. The maximum position shift is found to be 1.94 ± 0.03 mm (n=2), which is reduced to 1.17 ± 0.01 mm (n=3) and 0.98 ± 0.02 mm (n=4).

3.3.5 Contrast Analysis

A. Method

To test the contrast recovery performance of the proposed method, the same data used in the contrast analysis in the chapter 2 were used here. The activity concentration ratio between the lesion and the background is 10:1 and the acquisition time is set to 7 minute. Attenuation corrections for the breast phantom were made according to appendix A.4. Since no significant improvement is observed between 3 and 4 virtual DOI layers, only 3 layers were selected for comparison in the contrast study to save processing time. Two parameters were used for quantitative comparison: contrast recovery coefficient (CRC) and contrast-noise-ratio (CNR), as defined below. where c_{lesion} and a_{lesion} represent the reconstructed and real activity in the ROI of lesions. The method of ROI

selections is similar to that described in chapter 2.

$$CNR = \frac{c_{lession} - c_{background}}{\sigma_{background}} \tag{3.9}$$

$$CRC = \frac{(c_{lession} - c_{background})/c_{background}}{(a_{lession} - a_{background})/a_{background}} \tag{3.10}$$

B. Results

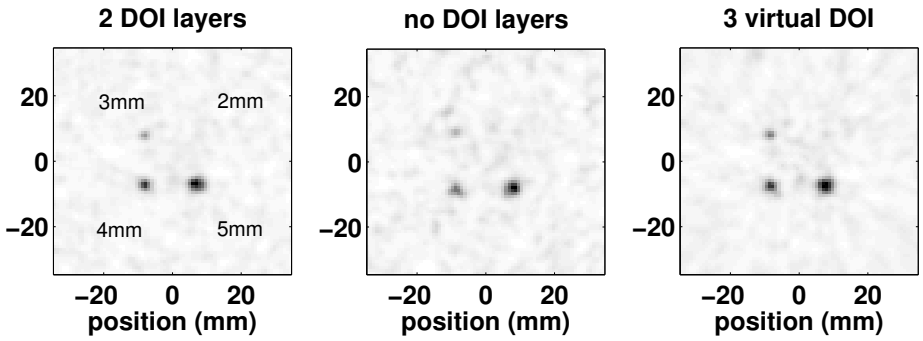

Figure 3.18: Reconstructed images of sphere lesions with activity concentration of 10:1 in the central slice, for system with/without DOI using shift invariant Gaussian kernel(left,middle) and resolution modeling method with 3 virtual DOI layer (right). Scan time in the simulation is 7 minutes. The images were reconstructed with 4 subsets and 8 iteration

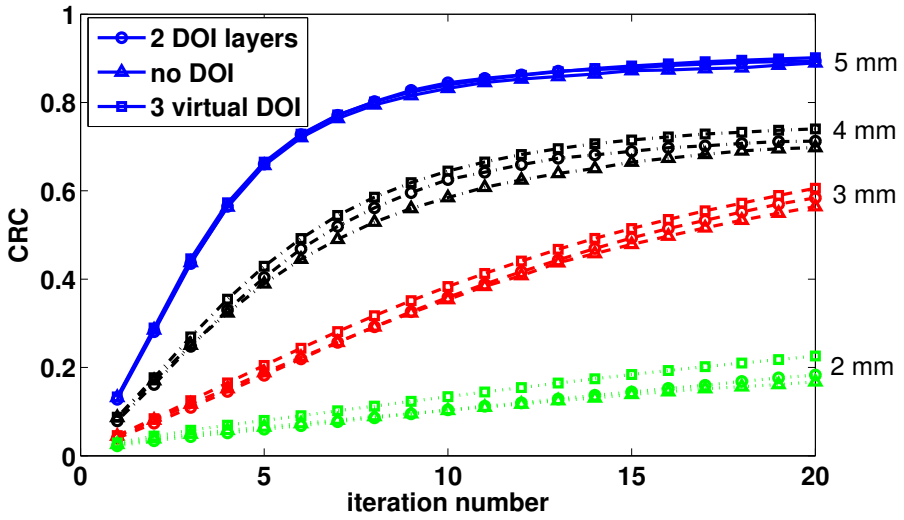

Figure 3.19: CRC results as a function of iteration number (4 subsets) for lesions with different sizes (5 mm, 4 mm, 3 mm and 2 mm).

Figure 3.18 shows the reconstructed images for three different cases (case 1, case 2 and case 3), using 4 subsets and 8 iterations. It is observed that the images of cases 2 and 3 exhibits clearer visualization of the lesions than case 1. As shown in Figure 3.19, the CRC results are plotted as a function of iteration number. For all lesions, CRC increases along with the increase of iteration numbers. Taking the 5 mm diameter lesion for an example, the CNR experience an initial steep increase and then gradually saturate, which is a well-known trend for convergence in OSEM implementation(Alessio and Kinahan, 2006). For smaller lesions (3 mm and 2mm), the CRC increases with a smaller slope due to slow convergence. On the other hand, the reason why larger lesions have higher CRC can be explained through modulation transfer function(Sorenson *et al.*, 1987). Consistent with the visualization, the cases 2 and 3 demonstrate higher CRC compared to case 1, particularly for spheres of 4 mm, 3 mm and 2 mm. For example, the CRC value for sphere 4 are 0.698, 0.713, and 0.740 for case 1, case 2, and case 3 when the iteration goes to 20. This result is consistent to the published relevant works(Pratx and Levin, 2011)(Zhang *et al.*, 2010).

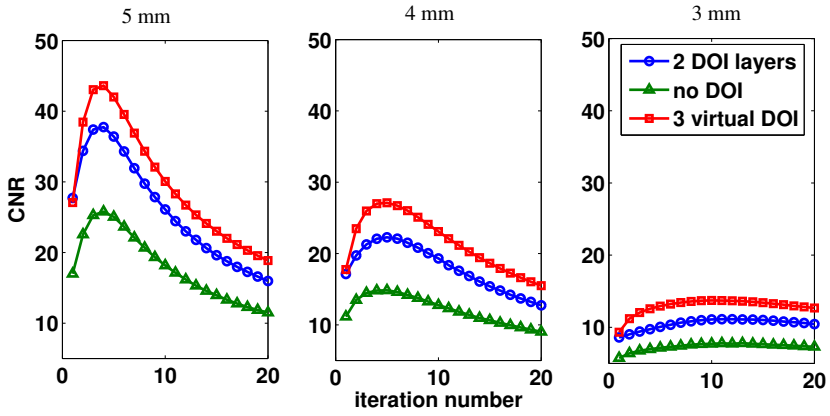

Figure 3.20: CNR results as a function of iteration number (4 subsets) for lesions with different sizes (5 mm, 4 mm, 3 mm and 2 mm).

CNR as a function of iteration number is shown in Figure 3.20 for three lesion sizes. Due to the difficulty in the selection of ROI as mentioned in chapter 2, the 2 mm diameter lesion was not included for CNR analysis. The results clearly indicate that case 3 is able to improve CNR for all three lesions and enhance detectability. For example, the peak CNR values for the 5mm sphere are 43.6, 37.7 and 25.8 for case 1, case 2 and case 3, respectively.

3.3.6 Consideration of processing speed and calculation time

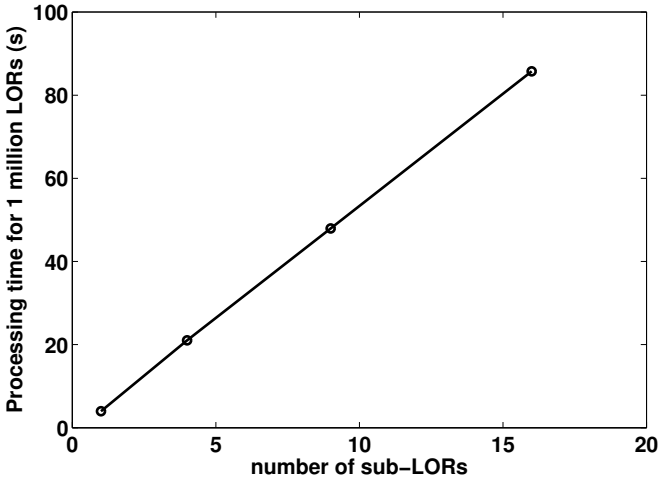

Figure 3.21: The processing time for 1 million LORs vs. the number of sub-LORs

Despite noticeable improvement on resolution and contrast found using the proposed method, processing speed and calculation time also need to be examined. For example, for the case 1 or 2 which is the simplest case, only an invariant spatial kernel is applied per LOR. However, for the proposed method, sub-LORs are generated and consequently the calculation time is to be increased depending on the selection of number of virtual DOI layers. Figure 3.21 shows the processing time required for implementing the OSEM with 1 iteration for 1 million LORs for different cases, which indicates that the processing time is proportional to the number of the sub-LORs created. For example, the processing time for case 3 using 9 sub-LOR (47.94s) is roughly 9/16 of that for case 3 using 16 sub-LOR (85.73s).

Although such workload increase may not pose a limitation on GPU implementation based on available commercial GPU processors, it is still desired to reduce the number of the LORs to be processed in order to speed up the reconstruction. To address this issue, we plan to: 1) convert the list-mode data to histogram data based on the combination of crystal IDs are grouped together

and processed together so only one round of back/forward projection is required; 2) dynamically select the number of virtual DOI layers based on geometry information of individual LORs. These solutions are currently being investigated in our group and only the concept is discussed below.

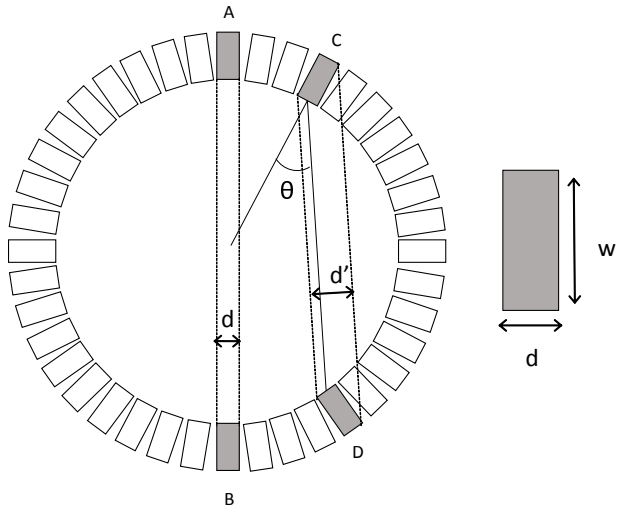

Figure 3.22: Variable number of virtual DOI layers

For example, as shown in Figure 3.22, a vertical LOR between two detectors in the center of FOV (facing each other and passing through the origin), no virtual DOI is necessary, whereas a vertical LOR at the edge of FOV suffers more from parallax error and thus requires a large number of virtual DOI layers. As shown in Figure 3.22, if we consider both A and B has DOI layers, then the LORs created by different DOI layers are identical in the image space. On the other hand, for the case of detector pair C and D, the width where the source could be detected is then not equal to the width of the crystal, it is given by:

$$d' = dcos\theta + wsin\theta \tag{3.11}$$

where d, θ and w are shown in the Figure 3.22. The width is significantly extended beyond the width of the crystal. Therefore, more virtual DOI layers are required and the differences among multiple sub-LORs are also expected to be more significant.

3.4 Discussion

The main motivation of this chapter is to develop a method to model the system matrix in OSEM reconstruction based on GPU platform to help us reduce the parallax error for the proposed breast PET/MR system and improve resolution. Our proposed method is based upon the concept of virtual DOI layers and is easy to implement with GPU. In addition, its flexibility allows it to be applied to other PET system geometries.

The critical question we intend to answer in this chapter is whether it is feasible to mitigate parallax error (i.e. DOI effect) for a PET system with the same accuracy as that obtainable with a physical DOI design. There exist several designs to incorporate DOI capability into PET detectors, for both scintillator array and semiconductor such as CZT (Peng and Levin, 2010). For instance, one design consists of two or more crystal layers each read out by a layer of photodetector(s)(Rafecas *et al.*, 2001). Another design employs a single crystal layer read out by two photodetectors on both ends, where the depth information of the annihilation photon inside the crystal is determined by the difference in the amount of light detected by the two photodetectors. Another common design is called phoswich design, which comprises two different types of scintillation crystal materials read out by the same photodetector(Mosset *et al.*, 2006). DOI capability is realized by pulse shape discrimination based on the fact that the decay time constants of two scintillation materials are different. A detailed review about DOI design can be found in (Peng and S Levin, 2010)

However, when any of these DOI designs is introduced, the cost of PET detector and system, as well as complexity is increased significantly. If a two-layer readout DOI design is used, the cost of photodetectors will be doubled. For some designs, the routine detector/system calibration is very challenging. It will be advantageous to use the proposed analytical method with GPU, if

comparable performances can be achieved.

The following section discusses four main aspects related to the proposed method: accuracy, spatial resolution, contrast and speed.

3.4.1 Accuracy

Our validation shows the comparison of results from our model against Monte-Carlo simulation and good agreement is observed (see Figure 3.5- 3.7). There are three limitations of our proposed model to be mentioned here, which are to be further studied in future work.

First, this model did not consider the inter crystal block penetration as illustrated in Figure 3.1, which was a physical process accurately modeled in GATE. This is one major source resulting in the minor discrepancy between the proposed method and Monte-Carlo simulation, which is most likely to affect LoRs on the edge. However, such implication allows us to directly calculate the system matrix on the fly with the GPU device, since calculating the attenuation length through adjacent crystal blocks may require much more intensive resources. A possible solution is to pre-calculate the intersection length offline on the CPU and save it as a look-up-table inside the GPUs fast cached memory, if necessary, with geometrical symmetry conditions taken into account to further reduce memory storage. Nevertheless, it should be emphasized that since no inter crystal block interaction was considered for all three cases we studied, the comparison is believed to be justified.

Second, the proposed model only considers sub-LORs that are created in different layers along the crystal depth position, while ignores the layers in the crystal surface directions. Since the crystal width is much smaller than the crystal thickness, there is minor difference to be expected(Qi *et al.*, 1998) between the calculated and the actual system matrix.

Third, the geometrical system matrix is based on a spatial invariant Gaussian kernel with the FWHM set to be equal to the crystal thickness (Pratx *et al.*, 2009). However, the FWHM of the Gaussian may deviate from the actual Gaussian kernel. For example, the width of a LOR that is created by two crystals with smaller offset will have larger FWHM value than that created by two crystals with bigger offset(Cui *et al.*, 2011a). This effect also needs to be further investigated.

3.4.2 Spatial Resolution

The proposed method is able to improve spatial resolution FWHM and reduce position shift, as shown in Figure 3.10- 3.17. Comparable performances between case 2 and case 3 gives us confidence that for a breast PET/MR system as well as other PET systems suffering from parallax error, the method based on virtual DOI layers we developed in this chapter is able to play an important role as an alternative solution to the solution of adding physical DOIs. Furthermore, it is found that once the number of virtual DOI layers exceeds 3, no significant improvement is observed (see Figure 3.16).

Another finding is that for both case 2 and case 3, the spatial resolution is still degrading towards the edge of the FOV (Figure 3.10), implying that the analytical calculated system matrix is still not able to represent the actual value. This effect is also found in a number of relevant published works (Strul *et al.*, 2003)(Stute *et al.*, 2011), and it is explained due to the exclusion of inter-crystal scatter modeling. In this work, it is also caused by the exclusion of inter crystal-block attenuation modeling. With carefully modeling of the system matrix through Monte Carlo or direct measurement, the spatial resolution is found to be uniform throughout the FOV(Panin *et al.*, 2006)(Stute *et al.*, 2011).

Compared to the case 2, the case 3 achieves better improvement on spatial resolution especially for the sphere located further away from the center. It is primarily because the system matrix for case 3 has better accuracy than case 2 near the edge. With the addition of two built-in DOI layers, the measured counts are better sampled, and dimension of the projection space is increased. Although

the system matrix is still calculated using the geometrical ray tracing method, the increased dimension in the projection space could help better solve the inverse problem to find the correct image. However, the geometrical system matrix is still based on a shift invariant Gaussian function, which has big deviation from the true value as the actual PSF is not purely Gaussian(Sureau *et al.*, 2008) especially near the edge. The deviation is even more obvious for the system without DOI design (see Figure 3.5), which caused the profile of the reconstructed point source to be shifted toward the edge and made it difficult to be quantified by Gaussian fitting (see Figure 3.9). Compared to case 2, the proposed method does not increase the dimension of the projection space, but better model the system matrix in which the number of the non-zero values in the system matrix is increased with improved accuracy. Therefore the spatial resolution at the edge of the FOV for the system with resolution modeling method is better than that using two built-in DOI layers.

3.4.3 Contrast

Based on the result (Figure 3.19), the proposed method improved the system contrast recovery performance. The improvement on contrast recovery could help in the identification and staging of tumors which is very important to the PET imaging for breast cancer. The possible reason for contrast recovery improvement is that the improved spatial resolution reduces the effect of partial volume effect (PVE), where smaller signal from the lesion is spreaded to the background(Soret *et al.*, 2007).

Furthermore, the noise level for the image obtained using resolution modeling method is lower resulting in higher CNR values shown in (Figure 3.20). This is due to the fact that the resolution modeling method modifies the noise structure(Rahmim *et al.*, 2013) which is intensively studied in (Tong *et al.*, 2010). A number of published works(Rahmim *et al.*, 2008)(Alessio *et al.*, 2010) which applied the resolution modeling method found that the convergence rate using the resolution modeling method is much lower than the conventional method. This is explained that more non-zero values are included in the system matrix so that the inverse problem of the image reconstruction needs more steps to solve. In addition, the resolution modeling method increases the

noise correlation between different voxels as more voxels are included in the forward and backward projection steps. (Tong *et al.*, 2010) derived equations found that the increase of noise correlation also contributes to the reduction of image noise determined by the spatial variance in a background ROI. Therefore the noise using the resolution modeling method could also be significantly reduced than the conventional method at the same iteration during the image reconstruction.

3.4.4 Speed

The resolution modeling method proposed in this work is applied to calculate the system matrix on the fly during the image reconstruction without saving it on the disk beforehand. Therefore the way to calculate the system matrix significantly affects the image reconstruction speed (see Figure 3.21). As found in this work, the processing time is proportional to the square of the number of the virtual DOI layers. However, it is also found that the spatial resolution improvement is more significant using more virtual DOI layers. Therefore, the trade off between the accuracy and the speed has to be balanced. Other than the two proposed solutions which are converting the list mode data to histogram data and using variable number of virtual DOI layers, there are also some other possible solutions which could reduce the processing time while retaining the accuracy. The first solution is to use more GPU devices to process the image reconstruction task. However, in order to efficiently use the computational power of all the GPU devices, the workload has to be distributed evenly. The second solution is to use geometrical symmetry to reduce the redundancy where a lot of system matrices have the same value. By doing this way, some of the system matrices have to be computed and saved in the memory on the GPU. Nonetheless, this solution could not save huge amount of data on the GPU and it is only helpful where the memory transfer time is much lower than the calculation time. Consequently, the possible new solutions need to be tested in the future.

Chapter 4

Conclusion and Future work

4.1 Conclusion

In conclusion, our group aims to develop a breast dedicated PET ring insert which could be integrated to the MRI bore to perform simultaneous PET/MR imaging. Monte Caro simulation was performed using GATE to evaluate the system and optimize the performance. The performance optimization is based on the result of photon sensitivity, NEC, spatial resolution and contrast recovery using different parameters such as crystal dimension, detector time/energy resolution and DOI resolution. However, not only the performance but also the design cost should be considered. The photon sensitivity was found to be proportional to the number of detector rings that is directly related to the biggest part of the design cost which is the number of scintillation crystals. Therefore, the final proposed system is configured to have an axial FOV of 33.6 mm, so that the resulting photon sensitivity (6.2%) is comparable to the existing breast dedicated PET devices. By carefully model the scanning condition using the XCAT human phantom, the NEC was found to be affected by time/energy resolution and time/energy window. With improved time and energy resolution, the NEC value gets higher. However the time and energy window needs to be carefully chosen. The best achievable NEC value was found to be 7886 cps with time resolution of 250ps, time window of 500ps energy resolution of 12% and energy window of 24%. The NEC was also found to be related

to the random events that are created by the contamination of singles counts emitted from the other parts of the body. With the addition of lead shielding, the NEC increased especially at high activity when the random events are increased. However at small activity setting (0.1μCi/cc) the improvement is not obviously seen. The spatial resolution was also tested. Based on the result, 2 DOI layers improved the spatial resolution and its uniformity, for the crystal with width of 2mm, with the DOI layer, the spatial resolution 50 mm away from center improved from 5.1±0.9 mm to 2.9±0.5 mm. Through the study of contrast recovery, TOF was found to be able to improve the contrast to noise ratio of the reconstructed images in a shorter scan time and fewer iteration numbers. Through this simulation, a better understanding of the parameters that determine the performance of the PET insert is achieved, which could help in the optimization of hardware design.

A resolution modeling method based on CDRF modeling and multiple ray tracing accelerated by the GPU device was also proposed aiming to reduce the parallax error caused by crystal penetration effect. The results showed that using the proposed method, the system is able to produce images with better spatial resolution with improved uniformity. In addition, it reduces the position offset between the object in the image and in the real space. It also improves the contrast recovery performance of the system and reduces the background noise. The processing time is found to be proportional to the number of the sub-LORs modeled. Several discussions on how to reduce the processing time were made. With the addition of resolution modeling method, it is expected to further improve the image quality for a prototyped system without upgrading the hardware. It also provides alternates to the design of the system DOI layer which contributes to the design cost and complexity. Therefore, for our proposed system, the system with resolution modeling is preferred to the DOI design.

4.2 Future work

Based on the simulation, a prototype system is going to be made and tested. The system also needs to be test with the integration of MRI device. Parameters setting are then based on the

Monte Carlo simulation result. The resolution modeling method and TOF method are going to be applied for image reconstruction. The resolution modeling method has a lot of space for further improvement in the future.

For example, the current method did not consider the inter crystal block attenuation which causes inaccuracies as shown in Figure 3.5. However, the calculation of the inter crystal block attenuation does not have the SIMD characteristic to be implement on GPU efficiently. Therefore, it is not included in this work, a better solution is to pre-calculate the crystal block attenuation on the CPU and save it as a look-up-table for reference.

Another limitation is that the processing time is much higher than the conventional method, and it is shown that this method is highly dependent on the number of the sub-LORs processed. To reduce the processing time, the number of sub-LORs could either be set to be different for each LOR or compressed using geometrical symmetry. In addition, multiple GPU devices could be used instead of one. Algorithm optimization needs to be done to ensure the work load difference for each processing unit is minimized.

The performance of spatial resolution improvement along the axial direction is not yet tested in this work, due to the small axial FOV of the proposed system. However this method is expected to provide better spatial resolution improvement for the system with larger axial FOV such as a dual-panel system.

The simultaneous incorporation of the resolution modeling method and the TOF method for image reconstruction is not yet tested. From the existing result obtained in this work, both methods increases the system contrast to noise ratio. Therefore the simultaneous incorporation of both method could further improve the image quality. In addition, for a dual-panel system (limit angle system), it is found that TOF reduces the spatial blurring caused by limit angle design(Surti and Karp, 2007). Consequently, the simultaneous incorporation of both methods is expected to further

improve the spatial resolution for limited angle system.

Appendix A

Appendix

A.1 Tomographic reconstruction

The conventional radiography approach applies the radiation to penetrate through the body and project onto a two dimensional plane to form an image. By referring the position and intensity distribution, the anatomical structures inside the body could be known. However, this image is the overall projection of many stacked slices of tissues, the individual anatomical information for each slice is missing. To get the 3D information, a 3D image is needed which could be produced by the tomographic approach.

The prefix "tomo" has the same meaning as "slice", therefore a tomogram is a slice of a 3D image. The idea of tomography is to collect projections from different view angles then reconstruct 3D image using tomographic reconstruction algorithms. The image could be reconstructed from detecting multiple emissions from outside the body such as CT or inside the body such as PET and SPECT, this work focus the PET image reconstruction where the emissions come from inside the body.

In PET, a positron annihilates with an electrons and forms two opposite 511 keV photons. When

two detectors detect the two photons in a coincidence window, we know that the annihilation occurs along the line connecting the two detectors, which is called a "line of response" (LOR). A line of response connects the projection information which is recorded by the detectors and the image information defined by the geometry. In terms of mathematical expression, the relationship between the projection and image could be written as:

$$r = xcos\phi + ysin\phi \tag{A.1}$$

$$s = ycos\phi - xsin\phi \tag{A.2}$$

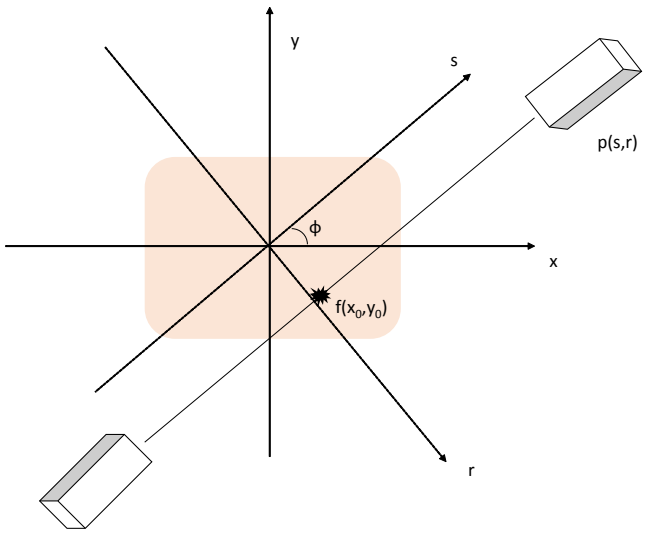

Figure A.1: The relationship between projection and image space

As shown in Figure A.1, the annihilation point f(x_0,y_0) in the 2D image space is detected by the detector pair defined by p(r,s). Other than the annihilation point, all the points f(x,y) in the 2D image along the line of response also follow equation A.1 and A.2. Therefore, given the projection information, all the points along the line of response could be the annihilation point, we

can backproject the line of response to all the image points given by equation A.1 and A.2. By changing the location of the detector and measuring the count, we can get a projection map for different r and ϕ, such map is called a sinogram as shown in Figure A.2 (B). The image is then reconstructed by backprojecting all the projection data recorded in the sinogram to the image space. However, as shown in Figure A.2 (C), the image reconstructed by simple backprojection has the blurring artifact compared to the true image A.2 (A). The blurring artifact is due to over-sampling the image in the center. To get an image with uniform sampling, the projection profile needs to be weighted before backprojection. A common approach is to take the Fourier transform of the projection profile and apply a filter to it. The filtered projection profile is then backprojected to the image, such process is called filtered backprojection (FBP). Compared to the image obtained from simple backprojection A.2 (C), the image A.2 (D) reconstructed by FBP has less blurring effect.

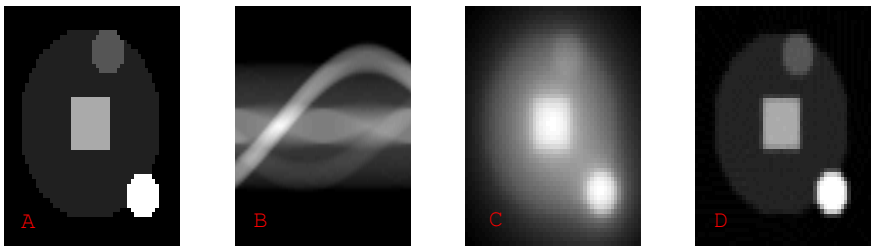

Figure A.2: Sinogram, simple backprojection and filtered backprojection

The non-uniformity of detection efficiency for the detectors could cause ring-shape blurring artifact in the final image reconstructed by the FBP (Figure A.3), which could be corrected by normalization correction. In the normalization correction, a uniform source is placed in the scanner as a reference to obtain the projection data, which could be divided by the uncorrected projection to obtain corrected projection data with uniform detection efficiency. However, in the projection space, there exist some dead areas where no detectors are available to detect the signals. Such area are not suitable for normalization by division. As shown in Figure A.4, for a ring shaped detector, the geometry configuration and sampling characteristic of two block detectors creates a diamond-like pattern in the sinogram. There exists some dead areas in the diamond which is due

to the non continuous sampling for a detector block pair. Moreover, there are dead areas between two diamond patterns, which corresponds to gaps between adjacent detector blocks. As there is no measured data, usually the dead areas are filled by interpolation, in this project linear interpolation is applied.

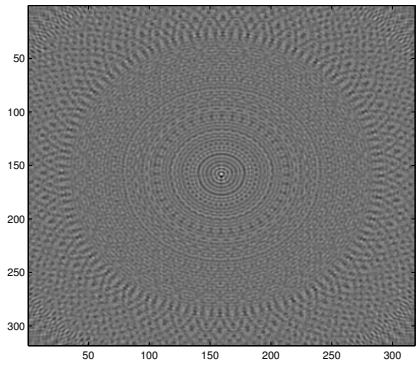

Figure A.3: The reconstructed image of a uniform source using FBP and Ramp filter without sinogram normalization correction. Ring-shaped artifact appears in the image.

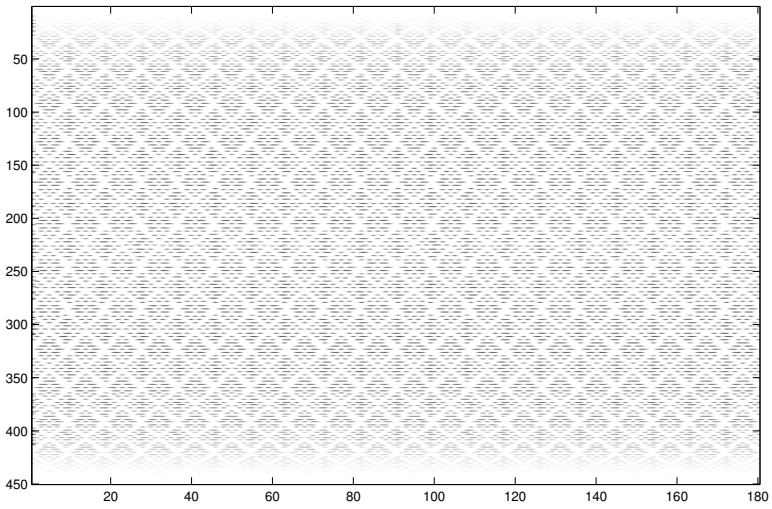

Figure A.4: The sinogram of a uniform source without normalization correction, diamond-like shape appears in the sinogram due to non-uniform detection efficiency.

In PET, other than placing detectors in a single plane which is described in the previous section, there are also other detectors set in different the axial positions. These detector pairs could not only create LORs in the direct planes, but could also create oblique LORs in the inter planes. As shown in Figure A.5, measuring only LORs in the direct planes is the 2D acquisition mode, measuring all the LORs including the oblique LORs is the 3D acquisition mode.

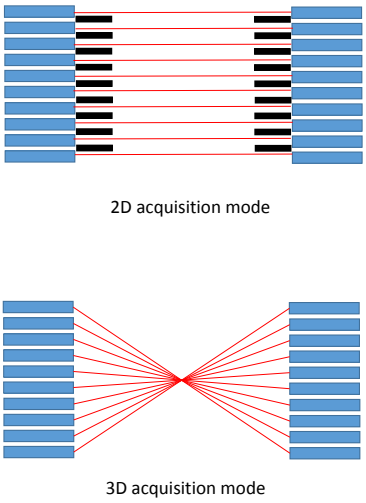

Figure A.5: PET acquisition modes, 2D (up), 3D (bottom)

In the 2D acquisition mode, the 3D image could be reconstructed by stacking slices of the 2D image reconstructed by 2D method. However, the septas between adjacent slices rejects the oblique LORs so that the detection efficiency is reduced which will consequently reduce the image SNR. To increase the detection efficiency, the speta could also be placed with a small slice offset to approximate that oblique LORs could be treated as in-plane LORs, but it still miss a significant amount of data. To fully use the available detectors to detect data, 3D acquisition mode is used where all the LORs are recorded, including those with a large ring difference. Directly using 2D reconstruction method for them could create spatial blurring in the axial direction, as they are backprojected in the wrong axial position. One solution to obtain 3D image using 3D data is to rebin the oblique LORs to a direct plane and such method is called rebining method. The simplest method is the Single Slice rebining method (SSRB)(Erlandsson *et al.*, 1994), where oblique LORs from two slices are rebined to the averaged slices. Another more accurate method is the Fourier rebining method(Defrise *et al.*, 1997), where the oblique sinograms are Fourier transformed and resorted to direct sinogram in the Fourier doman, and then transform back to to direct sinogram. Besides, the 3D data could be directly used for fully 3D image reconstruction, where no sinogram rebinning is required, but this method is more time consuming than the rebining methods.

A.2 Iterative reconstruction

An alternative to the FBP method (analytical reconstruction method) is known as the iterative reconstruction method. Compared to FBP, iterative reconstruction methods are more computationally intensive yet it could model the image noise and produce less noisy images. It is rarely applied in the early generation of PET scanners due to its mathematical complexity. However with the rapid evolution of computer speed, they are more often used in the current PET systems. The concept of iterative reconstruction could be summarized as shown in Figure A.6. The first step is to initialize the image space, usually the image is set to have uniform value. After initialization, image is forward projected to get a calculated projection value, this process is called forward projection. The calculated projection value is then compared with the measured value. After that, the image is updated by backprojection to make sure the calculated projection is equal to the measured projection value. This process is then looped again with the updated image until the difference between the calculated projection value for the estimated image and the measured projection value is less than a predefined value.

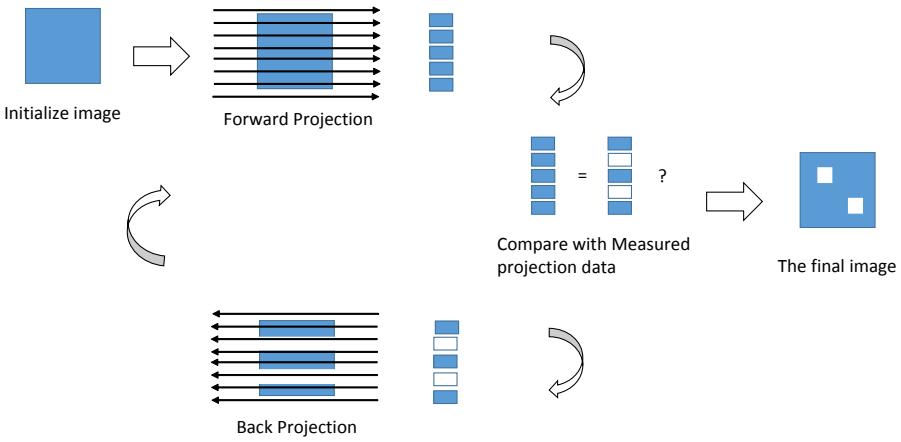

Figure A.6: The iterative reconstruction process

A common iterative method for PET is the Maximum Likelihood Expectation Maximization (MLEM) method(Dempster *et al.*, 1977) (Shepp and Vardi, 1982), whose noise model is based on the Poisson

distribution of the photon detection. The likelihood of of detecting P counts which is equal to the true count p, given the image x is defined by:

$$L(P = p|x) = \prod_{i=1} \frac{\bar{p}_i^{p_i} exp(-(\bar{p}_i)}{p_i!} \tag{A.3}$$

Here, $\bar{p}$ could also be expressed as $\bar{p} = Hx$ where H is the system matrix connecting the projection vector $\bar{p}$ and the image vector x. Taking the logarithm of the likelihood and maximize it, the MLEM equation could then be written as:

$$x_j^{n+1} = \frac{x_j^n}{N_j} \sum_{i=1} H_{ij} \frac{p_i}{\sum_{j=1}^N H_{ij} x_j^n} \tag{A.4}$$

where i and j are the IDs of the LORs and image, p is the measured projection number, H is the system matrix and x is the image. The image x^n at n^{th} iteration is updated in the iteration n to get the updated image x^{n+1} for the (n+1)th iteration. The term $\sum_{j=1}^N p_{ij} x_j^n$ represents the forward projection process. $\sum_{i=1}^P H_{ij} \frac{p_i}{\sum_{j=1}^N H_{ij} x_j^n}$ stands for the backprojection, and N_j is the normalization term.

The MLEM algorithm iteratively compares the projection and updates the image until the image reach acceptable value. Such process could be very time consuming because usually more than 50 iterations are required. An improved solution to MLEM, the Ordered Subset Expectation Maximization (OSEM) (Hudson and Larkin, 1994) was then proposed to speed up the reconstruction process. OSEM is an extension to MLEM, where instead of updating image by processing the whole LOR data set, subsets of all LORs were used to update image. The process could could be expressed in the following equation:

$$x_j^{n+1} = \frac{x_j^n}{N_j} \sum_{i \subset S_b} H_{ij} \frac{p_i}{\sum_{j=1}^N H_{ij} x_j^n} \tag{A.5}$$

The term S_b corresponds to the number of LORs within a subset, when there is only one subset, OSEM is equivalent to MLEM.

Both equation A.4 and A.5 applies to the case where the data are recorded in the sinogram format. However, in some cases, individual coincidence data are recorded directly into list mode without histograming to sinogram. The benefit of the list mode acquisition mode includes several reasons. First, the sinogram is predefined to store data recorded in a wide range of angle and radial offset. In many cases, the sinogram is very sparse that the counts are detected in a limited range of angle and radial offset. For list mode data, the zero projection are not considered in the image reconstruction process. Consequently, the system matrix for zero projection is not required to be calculated, this allows the system to apply on-the-fly calculation of system matrix. Second, the list mode data consider each LOR individually, which could be useful to apply the time of flight reconstruction method as the time differences for LORs detected by two detectors are not the same. Finally, the list mode data represent the raw LOR without any spatial blurring and correction as is done in the sinogram. In terms list mode data, the OSEM equation could be written as:

$$x_j^{n+1} = \frac{x_j^n}{N_j} \sum_{i \subset S_b} H_{ij} \frac{1}{\sum_{j=1}^{N} H_{ij} x_j^n} \tag{A.6}$$

A.3 GPU accelerated OSEM

The modern medical imaging devices are producing larger amount of data, and are requiring more advanced reconstruction methods with larger computational burden. Moore's law shows that the performance of computer doubles approximate 18 months, however it could not meet the increasing demand for fast growing image reconstruction and processing work. With the help of parallel computing, the whole reconstruction task could be sub-divided into many small tasks, which could be processed together using different computers or on a cluster computer. In this way, the whole processing time is divided by the number of computer or the number of the nodes on the cluster. However, the space multiple computer or cluster computer requires is huge, and special cooling systems are required. In addition, more electrical power is needed to keep the system running.

This creates difficulties for a hospital allocate a space to locate the whole system. The recent development of general purpose Graphic Processing Unit has enlightened people to migrate the image reconstruction task from CPU to the GPU.

Graphic Processing Unit (GPU), is initially dedicated to accelerate and manipulate the computer graphics tasks which are formerly processed by the CPU. Compared to CPU, it has many processing units which could process the transforming and rendering of the small geometrical elements such as triangles at the same time. The rapid development of computer game and computer aided design had increased the demand of advance GPUs. New generation of GPUs are developed that made programmers to do more complex jobs, and recent GPUs have included the general purpose computing capability, where instead of computer graphic work, more general computing on the CPU could be migrated to the GPU. Because of the large number of processing units and the high memory bandwidth, GPU has superior advantage over CPU when dealing with large amount of data that need to be processed using similar instruction, such as matrix operations which is common in the image reconstruction and processing. Dedicated softwares such as OpenCL and CUDA have been developed for programmers to interface with the GPU easily using analogous programming format of the C language. Though OpenCL is also a powerful software for GPGPU development, this work only focus on the development using the CUDA.

CUDA is developed by the Nvidia corporation, which is a programming interface optimized for Nvidia GPUs. It utilized the single instruction multiple data (SIMD) characteristic of the GPGPU computing, where each processing unit is computing the same instruction, which is defined in a specific function called a kernel, with different data input. CUDA has defined a structure to group many processing units, in addition it has specific memory architecture to optimize the memory transfer rate for different types of data.

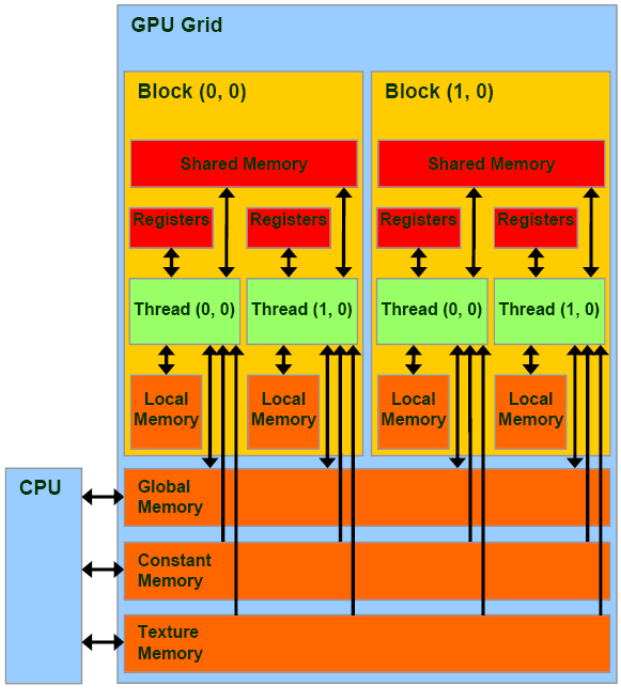

Figure A.7: The CUDA architecture (Nvidia, 2011)

As shown in Figure A.7(Nvidia, 2011), the CUDA architecture has defined 4 types of memory, global memory, constant memory, texture memory and shared memory, all are located on the GPU. The global memory has the largest memory space, allowing users to transfer data from CPU to the GPU through the PCI interface. The data transfer rate within the global memory is faster than the transfer rate between CPU and GPU through the PCI, but it is slower than the other 3 types of memory. Therefore the global memory is usually used to store raw data on the GPU which are processed by the processing units. CUDA defined the smallest processing units as threads, which are grouped into blocks, and blocks are grouped into grids. Shared memory is defined in each blocks, where the threads could access the shared memory quickly only within the same block. The shared memory is optimized to store temporary value which needs to be accessed and updated by different threads constantly. The other two types of memory, constant memory and texture memory are both read-only memory, they could be accessed quickly by all the threads lunching the kernel,

the texture memory has built-in interpolation function which is optimized for image processing. In this work, only global and shared memory are used.

A GPU based fully 3D list mode OSEM program is developed in this project which implements the computing model proposed by (Cui *et al.*, 2011b), since it utilizes the GPU performance very efficiently. Several extensions to this model are made which are described in chapter 3. In this model, each LOR is processed by a thread independently and simultaneously. A tube-of-response method is applied to calculate the system matrix used for the list-mode OSEM.

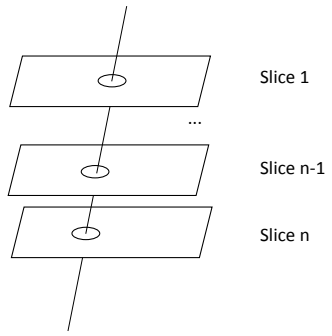

Figure A.8: Calculation of system matrix in the 3D image space

As shown in Figure A.8, instead of projecting the LORs directly in the 3D image space, the 3D image is first sliced to different 2D layers, each layer has an intersection point of the LOR passing through it. By calculating the distance from the voxel in the slice to the intersection point of the LOR, the system matrix is obtained by the following Gaussian kernel:

$$H_{ij} = exp(\frac{-d_{ij}^2}{2\sigma^2}) \tag{A.7}$$

where d_{ij} is the difference between voxel j and the LOR i, σ is selected to control the blurring effect of the the Gaussian function, which in this work is chosen to have the FWHM value equal to the crystal width. Instead of calculating the distance from the intersection point for all the voxels in the slice, a boundary is set to cut-off the pixels that is very far away from the LOR. However, this

posses a problem selecting the boundary, for the LORs perpendicular to the slice, the boundary on the slice is much smaller than the LORs having a small intersection angle to the slice, even though the voxels on the boundary have the same distance to the LOR for both cases, which could be illustrated in the Figure A.9.

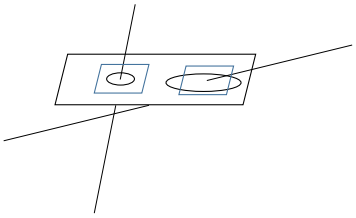

Figure A.9: Tube of response boundary selection

Defining different boundary size causes the problem that the work load for each threads is different, since the threads are processing simultaneously, some threads may have finished all the work waiting for other threads to finish, this significantly limit the processing speed. A solution is to reorient the image to make sure the intersection angle between the slice and LORs have a larger value, where in this model the value is set to be larger than 45 degree, then the same boundary is selected to effectively cut-off the pixels far away from the LORs for all the LORs.

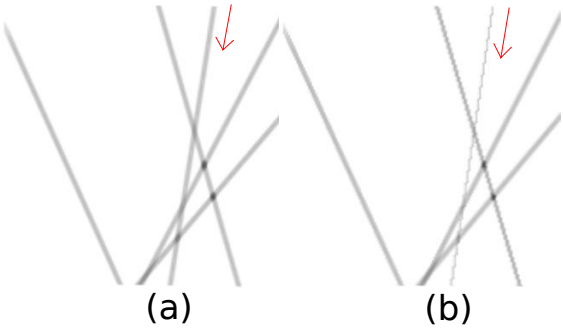

Figure A.10: The effect of choosing the same boundary without reorienting the LORs

In this work, LORs are coded as the ID of the crystals which is stored as 16-bit integer on the global memory, which could be decoded to calculate the x,y,z positions of two crystals that form a LOR.

The image is defined in the shared memory, where each thread could read and update its value very efficiently while processing the forward and backprojection. The complete reconstruction process could be summarized as the following process. First, the LORs are separated to three groups, each has the largest intersection angle for one of the xy, xz and yz plane of the 3D image. Second, each LORs are forward projected to the image space by projecting to its corresponding consecutive slices, the forward projection value is accumulated and stored in the global memory. Then each LORs are backprojected to the image space again updating the image value slice by slice. Finally, the image is divided by the normalization image and transferred to the next loop.

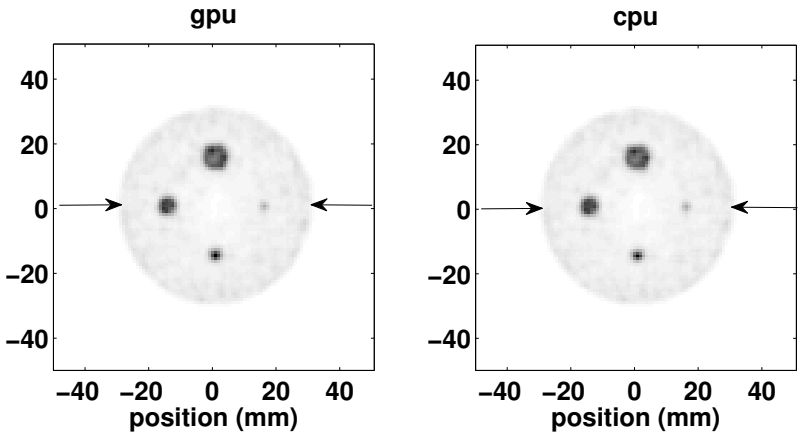

Figure A.11: Comparison of CPU (right) and GPU (left) reconstructed images using list-mode OSEM with 4 subsets and 8 iterations for 4 million LORs.

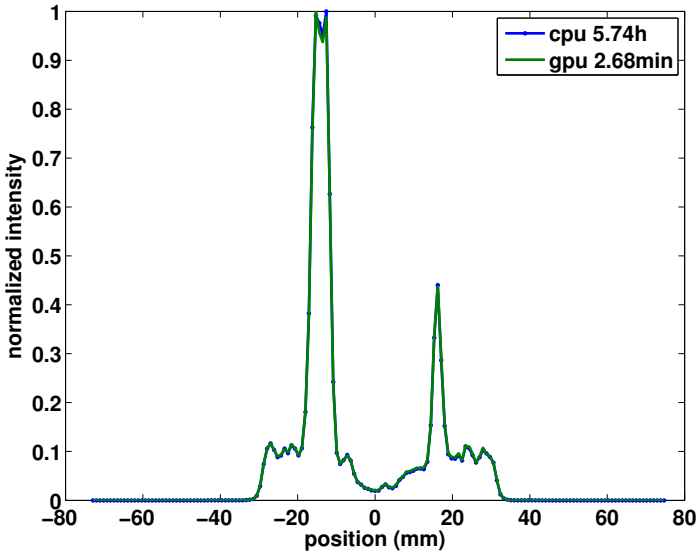

Figure A.12: Line profile of image reconstructed with CPU and GPU shown in Figure A.11. CPU takes 5.74h and GPU takes 2.68 min in the reconstruction.

Figure A.11 shows the image reconstructed with list-mode OSEM on both CPU (Intel E5700) and GPU (Nvidia GTX 465). Performing the same 4 million LORs data with 4 subsets and 8 iterations, GPU takes 2.68 min while CPU needs 5.74h. From the result shown in Figure A.12, the line profiles created along the same line on the image reconstructed by CPU and GPU appear identical.

A.4 Attenuation correction

When two 511 keV photons are created by annihilation, it will interact with the tissue before detected by the detectors. The energy of the photon may fall below the energy window of the detector through photoelectric effect or Compton scattering in the tissue causing the data missing. In addition, the trajectory of the photon may alter through Compton scattering or coherent scattering causing the data missing in which the photons may escape from the FOV. In terms of mathematical expression, the number of photons N transmitted through an object with attenuation coefficient μ

is found to be:

$$N = N_0 exp(-\int_s \mu(x,y)dr) \tag{A.8}$$

where N_0 is the number of photons created that are transmitting through the tissue and μ is the attenuation coefficient. For PET, as there are two photons created along opposite directions, the integral in equation A.8 represents the total path where the two photons transmit through the tissue. However, the attenuation coefficient for the photons along the transmission path is not always uniform. Hence an accurate distributions of attenuation coefficient in the body is required for the attenuation correction which is called a attenuation map(Zaidi and Hasegawa, 2003).

One of the approach to obtain the attenuation map is carried out by the CT scan. Almost all of the clinical whole body PET systems are now integrated with the CT component, which produces images that reflects the different attenuation distributions in the body. Another method is to calculate the attenuation map by segmentation of the image such as MRI image to find different types of tissue in the body.

In this project, where the tissue to be imaged is breast where the attenuation coefficient distributions could be treated as uniform. Therefore, to correct the effect of attenuation, only the transmitting length for the two 511 keV photons need be calculated.

To calculated the transmitting length, a simplified geometry condition is applied in which the breast could be treated as a cylinder because of the small axial FOV of the PET ring detector. For a LOR detected by two detector located at (x_0, y_0, z_0) and (x_1, y_1, z_1), the transmitting length

L_{atten} could be calculated using the following equation:

$$L_{atten} = 2\sqrt{R^2 - L^2}/cos\theta \tag{A.9a}$$

$$L = |(-x_0, -y_0)| \times sin\phi \tag{A.9b}$$

$$cos\phi = \frac{(-x_0, -y_0) \cdot (x_1 - x_0, y_1 - y_0)}{|(-x_0, -y_0)| \times |(x_1 - x_0, y_1 - y_0)|} \tag{A.9c}$$

$$sin\theta = |z_0 - z_1|/|(x_0 - x_1, y_0 - y_1, z_0 - z_1))| \tag{A.9d}$$

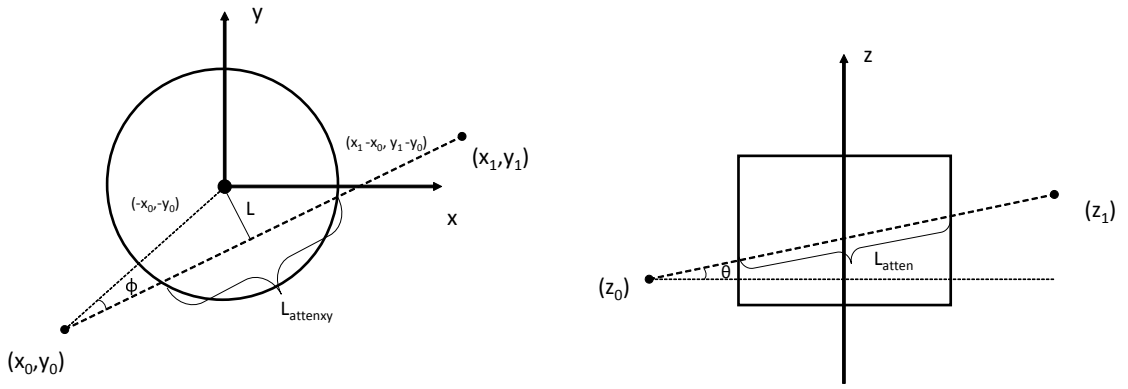

Figure A.13: Attenuation length calculation for attenuation correction

With the attenuation correction, the number of projection is corrected, therefore the OSEM equation could be written as:

$$x_j^{n+1} = \frac{x_j^n}{N_j} \sum_{i \subset S_b} H_{ij} \frac{p_i / e^{-\mu L_a tten}}{\sum_{j=1}^{N} H_{ij} x_j^n} \tag{A.10}$$

Figure A.14 shows the reconstructed image of cylinder source filled with water with a radius of 3 cm with and without attenuation correction. The image is reconstructed by list-mode OSEM with 2 subsets and 4 iterations.

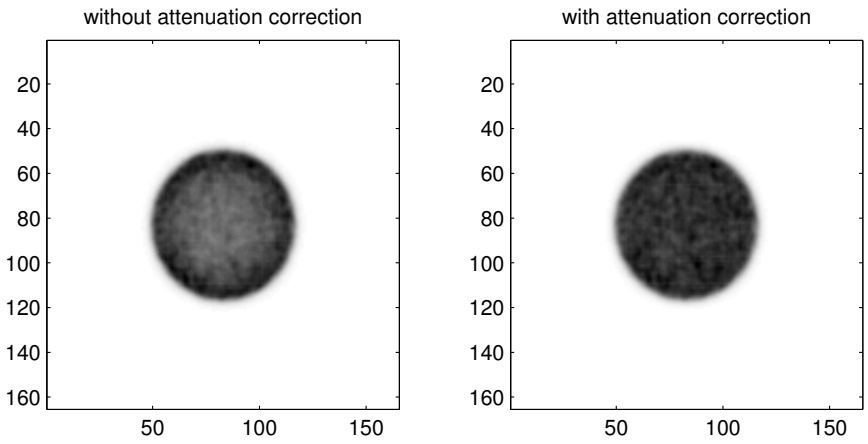

Figure A.14: Uniform phantom filled with water reconstructed with/without attenuation correction

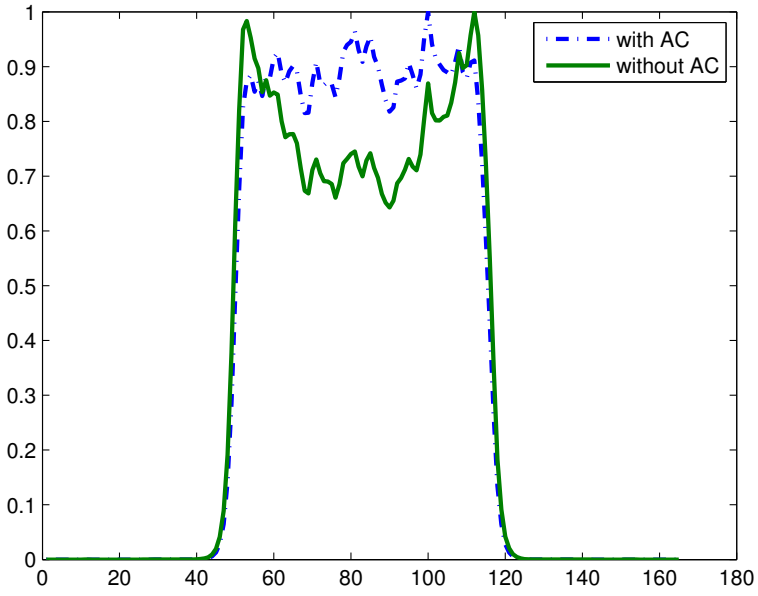

Figure A.15: Line profile of the uniform phantom reconstructed with/without attenuation correction

As shown in Figure A.15, for the image obtained without attenuation correction, the intensity at the

center of the cylinder is smaller than that at the surface of the cylinder. However, the intensity for the image reconstructed with attenuation correction is more uniform distributed from the surface to the center.

A.5 Time of flight image reconstruction

with the incorporation of time of flight function, the PET scanner could detect the time difference with improved accuracy which could be used to locate the annihilation point.

As shown in Figure A.16, if the time at which the 511 keV photons are detected are t_1 and t_2, then the difference from the center of the LOR for the annihilation point could be calculated as:

$$\Delta x = \frac{c}{2}(t_1 - t_2) \tag{A.11}$$

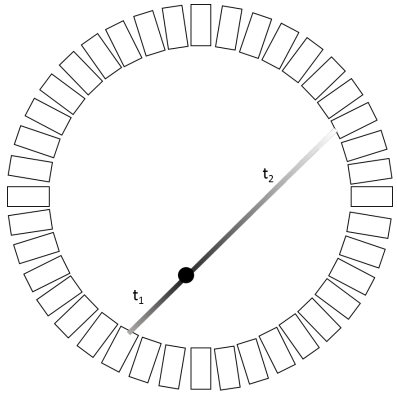

Figure A.16: Illustration of time of flight concept

However, there is no such a perfect system to measure the exact time difference. The accuracy of the time difference depends on the system's time resolution. According to equation A.12, the probability distributions of the annihilation point follows the probability of the time difference, which could be used to assign the position of the annihilation point with a weight. As shown in Figure A.16, the

exact annihilation point is represent as a dot, in the conventional image reconstruction process, all the points LOR passes have the same probability of being detected, whereas in the case of TOF, the probability of detection is different which is given as:

$$H_{tof} = e^{\frac{(t_1 - t_2)^2}{2\sigma_t}} \tag{A.12}$$

$$H_{tof} = e^{\frac{\Delta x}{2\sigma_x}} \tag{A.13}$$

where σ_t is determined by the system time resolution. The annihilation position along the LOR calculated using equation A.9 has a probability of H_{tof} given the time difference (t_1-t_2).

Since each LOR has a unique time difference, usually the TOF method is applied to the list mode data, where each LOR is processed independently. In this project, the TOF method is applied to the list-mode OSEM, where the system matrix is the product of the geometrical term and the TOF term.

$$H = H_{geom} H_{tof} \tag{A.14}$$

As shown in the equation A.13, the probability distribution of a source located along the LOR is Gaussian, where probability of the source that has a distance greater than $3\sigma_x$ from the LOR center is approximated to be zero. Therefore, the length of the LOR and the number of the system matrix needed to be calculated could be reduced in order to speed up the reconstruction.

In order to efficiently utilize the power of GPU's SIMD architecture, the work load of each processing unit should be adjusted to be balanced. However, processing in the conventional way described in section A.3 where each slice of the 3D volume is processed sequentially causes the problem that for LOR with different time difference, the work load on each slice is different, because the sources on the slice may have distance greater than $3\sigma_x$ to the center. In this case, the LORs are grouped according to the time difference before the image reconstruction, so that the work load for the LOR

within the group is balanced. The sorting is processed by the Nvidia's Thrust Library which takes milliseconds to process millions of LORs.

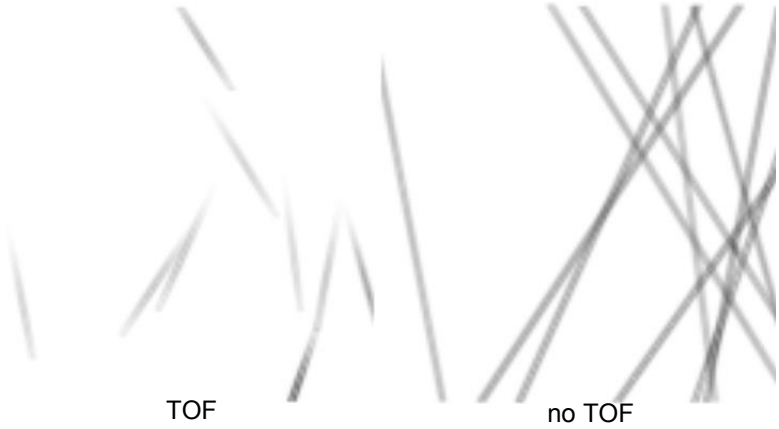

Figure A.17: The effect of projection using TOF method (left) and the conventional method (right).

Figure A.17 shows the effect of TOF for a system with 250ps time resolution by backprojecting 10 LORs, as shown in the figure, the length of the LOR using TOF method is reduced and the position of the LOR is distributed sparsely and randomly based on the time difference.

Bibliography

Abreu, M., Aguiar, D., Albuquerque, E., Almeida, F., Almeida, P., Amaral, P., Auffray, E., Bento, P., Bruyndonckx, P., Bugalho, R., *et al.* (2007). Clear-pem: A pet imaging system dedicated to breast cancer diagnostics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **571**(1), 81–84.

Agostinelli, S., Allison, J., Amako, K. e., Apostolakis, J., Araujo, H., Arce, P., Asai, M., Axen, D., Banerjee, S., Barrand, G., *et al.* (2003). Geant4a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **506**(3), 250–303.

Aguiar, P., Rafecas, M., Ortuño, J. E., Kontaxakis, G., Santos, A., Pavía, J., and Ros, D. (2010). Geometrical and monte carlo projectors in 3d pet reconstruction. *Medical physics*, **37**, 5691.

Alessio, A. and Kinahan, P. (2006). Pet image reconstruction. *Nuclear Medicine*, **1**.

Alessio, A. M., Kinahan, P. E., and Lewellen, T. K. (2006). Modeling and incorporation of system response functions in 3-d whole body pet. *Medical Imaging, IEEE Transactions on*, **25**(7), 828–837.

Alessio, A. M., Stearns, C. W., Tong, S., Ross, S. G., Kohlmyer, S., Ganin, A., and Kinahan, P. E. (2010). Application and evaluation of a measured spatially variant system model for pet image reconstruction. *Medical Imaging, IEEE Transactions on*, **29**(3), 938–949.

Baro, J., Sempau, J., Fernandez-Varea, J., and Salvat, F. (1995). Penelope: an algorithm for monte carlo simulation of the penetration and energy loss of electrons and positrons in matter. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, **100**(1), 31–46.

Briesmeister, J. F. (1986). Mcnp-a general monte carlo code for neutron and photon transport. *LA-7396-M*.

Buchbender, C., Heusner, T. A., Lauenstein, T. C., Bockisch, A., and Antoch, G. (2012). Oncologic pet/mri, part 1: tumors of the brain, head and neck, chest, abdomen, and pelvis. *Journal of Nuclear Medicine*, **53**(6), 928–938.

Budinger, T. F. (1983). Time-of-flight positron emission tomography: status relative to conventional pet. *Journal of nuclear medicine*, **24**(1), 73–78.

Budinger, T. F. (1998). Pet instrumentation: what are the limits? In *Seminars in nuclear medicine*, volume 28, pages 247–267. Elsevier.

Buvat, I., Castiglioni, I., *et al.* (2002). Monte carlo simulations in spet and pet. *QJ Nucl. Med*, **46**(1), 48–61.

Camarda, M., Belcari, N., Guerra, A. D., Galeotti, S., Morsani, F., Herbert, D. J., and Vaiano, A. (2006). Development of the yap-pem scanner for breast cancer imaging. *Physica Medica*, **21**, 114–116.

Castiglioni, I., Cremonesi, O., Gilardi, M., Bettinardi, V., Rizzo, G., Savi, A., Bellotti, E., and Fazio, F. (1999). Scatter correction techniques in 3d pet: a monte carlo evaluation. *Nuclear Science, IEEE Transactions on*, **46**(6), 2053–2058.

Catana, C., Wu, Y., Judenhofer, M. S., Qi, J., Pichler, B. J., and Cherry, S. R. (2006). Simultaneous acquisition of multislice pet and mr images: initial results with a mr-compatible pet scanner. *Journal of Nuclear Medicine*, **47**(12), 1968–1976.

Catana, C., Procissi, D., Wu, Y., Judenhofer, M. S., Qi, J., Pichler, B. J., Jacobs, R. E., and Cherry, S. R. (2008). Simultaneous in vivo positron emission tomography and magnetic resonance imaging. *Proceedings of the National Academy of Sciences*, **105**(10), 3705–3710.

Conti, M. (2009). State of the art and challenges of time-of-flight pet. *Physica Medica*, **25**(1), 1–11.

Cui, J., Pratx, G., Prevrhal, S., Zhang, B., Shao, L., and Levin, C. S. (2011a). Measurement-based spatially-varying point spread function for list-mode pet reconstruction on gpu. In *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pages 2593–2596. IEEE.

Cui, J.-y., Pratx, G., Prevrhal, S., and Levin, C. S. (2011b). Fully 3d list-mode time-of-flight pet image reconstruction on gpus using cuda. *Medical physics*, **38**(12), 6775.

Daube-Witherspoon, M. E. and Carson, R. E. (1991). Unified deadtime correction model for pet. *Medical Imaging, IEEE Transactions on*, **10**(3), 267–275.

Daube-Witherspoon, M. E. and Muehllehner, G. (1987). Treatment of axial data in three-dimensional pet. *J Nucl Med*, **28**(11), 1717–1724.

de Jong, H. W., Boellaard, R., Knoess, C., Lenox, M., Michel, C., Casey, M., and Lammertsma, A. A. (2003). Correction methods for missing data in sinograms of the hrrt pet scanner. *Nuclear Science, IEEE Transactions on*, **50**(5), 1452–1456.

Defrise, M., Kinahan, P. E., Townsend, D. W., Michel, C., Sibomana, M., and Newport, D. (1997). Exact and approximate rebinning algorithms for 3-d pet data. *Medical Imaging, IEEE Transactions on*, **16**(2), 145–158.

Delorme, S. and Knopp, M. (1998). Non-invasive vascular imaging: assessing tumour vascularity. *European radiology*, **8**(4), 517–527.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 1–38.

Doll, J., Oberdorfer, F., and Lorenz, W. J. (1997). Performance evaluation of a whole-body pet scanner using the nema protocol. *J Nucl Med*, **38**, 1614–1623.

Downie, E., Yang, X., and Peng, H. (2013). Investigation of analog charge multiplexing schemes for sipm based pet block detectors. *Physics in medicine and biology*, **58**(11), 3943.

Eklund, A., Dufort, P., Forsberg, D., and LaConte, S. M. (2013). Medical image processing on the gpu-past, present and future. *Medical Image Analysis*.

Erlandsson, K., Esser, P., Strand, S., and Van Heertum, R. (1994). 3d reconstruction for a multi-ring pet scanner by single-slice rebinning and axial deconvolution. *Physics in medicine and biology*, **39**(3), 619.

Fischer, U., Kopka, L., and Grabbe, E. (1999). Breast carcinoma: Effect of preoperative contrast-enhanced mr imaging on the therapeutic approach1. *Radiology*, **213**(3), 881–888.

Furuta, M., Kitamura, K., Ohi, J., Tonami, H., Yamada, Y., Furumiya, T., Satoh, M., Tsuda, T., Nakazawa, M., Hashizume, N., *et al.* (2009a). Basic evaluation of a c-shaped breast pet scanner. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pages 2548–2552. IEEE.

Furuta, M., Kitamura, K., Ohi, J., Tonami, H., Yamada, Y., Furumiya, T., Satoh, M., Tsuda, T., Nakazawa, M., Hashizume, N., *et al.* (2009b). Basic evaluation of a c-shaped breast pet scanner. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pages 2548–2552. IEEE.

Gambhir, S. S. (2002). Molecular imaging of cancer with positron emission tomography. *Nature Reviews Cancer*, **2**(9), 683–693.

Goerres, G. W., Michel, S. C., Fehr, M. K., Kaim, A. H., Steinert, H. C., Seifert, B., von Schulthess, G. K., and Kubik-Huch, R. A. (2003). Follow-up of women with breast cancer: comparison between mri and fdg pet. *European radiology*, **13**(7), 1635–1644.

Gu, Y., Pratx, G., Lau, F. W., and Levin, C. S. (2010). Effects of multiple-interaction photon events in a high-resolution pet system that uses 3-d positioning detectors. *Medical physics*, **37**, 5494.

Habte, F., Foudray, A., Olcott, P., and Levin, C. (2007). Effects of system geometry and other physical factors on photon sensitivity of high-resolution positron emission tomography. *Physics in medicine and biology*, **52**(13), 3753.

Herbert, D., Moehrs, S., Dascenzo, N., Belcari, N., Del Guerra, A., Morsani, F., and Saveliev, V. (2007). The silicon photomultiplier for application to high-resolution positron emission tomography. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **573**(1), 84–87.

Hudson, H. M. and Larkin, R. S. (1994). Accelerated image reconstruction using ordered subsets of projection data. *Medical Imaging, IEEE Transactions on*, **13**(4), 601–609.

Jan, S., Santin, G., Strul, D., Staelens, S., Assie, K., Autret, D., Avner, S., Barbier, R., Bardies, M., Bloomfield, P., *et al.* (2004). Gate: a simulation toolkit for pet and spect. *Physics in medicine and biology*, **49**(19), 4543.

Judenhofer, M. S., Catana, C., Swann, K., Siegel, S., Jung, W., Nutt, R., Cherry, S., Claussen, C., and Pichler, B. (2006). Simultaneous pet/mr images, acquired with a compact mri compatible pet detector in a 7 tesla magnet. *Radiology*.

Judenhofer, M. S., Wehrl, H. F., Newport, D. F., Catana, C., Siegel, S. B., Becker, M., Thielscher, A., Kneilling, M., Lichy, M. P., Eichner, M., *et al.* (2008). Simultaneous pet-mri: a new approach for functional and morphological imaging. *Nature medicine*, **14**(4), 459–465.

Kalinyak, J. E., Schilling, K., Berg, W. A., Narayanan, D., Mayberry, J. P., Rai, R., DuPree, E. B., Shusterman, D. K., Gittleman, M. A., Luo, W., *et al.* (2011). Pet-guided breast biopsy. *The breast journal*, **17**(2), 143–151.

Karp, J. S., Surti, S., Daube-Witherspoon, M. E., and Muehllehner, G. (2008). Benefit of time-of-flight in pet: experimental and clinical results. *Journal of Nuclear Medicine*, **49**(3), 462–470.

Khalkhali, I., Mena, I., and Diggles, L. (1994). Review of imaging techniques for the diagnosis of breast cancer: a new role of prone scintimammography using technetium-99m sestamibi. *European journal of nuclear medicine*, **21**(4), 357–362.

Kim, C. L., Wang, G.-C., and Dolinsky, S. (2009). Multi-pixel photon counters for tof pet detector and its challenges. *Nuclear Science, IEEE Transactions on*, **56**(5), 2580–2585.

Lecomte, R., Schmitt, D., and Lamoureux, G. (1984). Geometry study of a high resolution pet detection system using small detectors. *Nuclear Science, IEEE Transactions on*, **31**(1), 556–561.

Levin, C. S. and Hoffman, E. J. (1999). Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution. *Physics in medicine and biology*, **44**(3), 781.

Liang, Z. (1994). Detector response restoration in image reconstruction of high resolution positron emission tomography. *Medical Imaging, IEEE Transactions on*, **13**(2), 314–321.

MacDonald, L., Edwards, J., Lewellen, T., Haseley, D., Rogers, J., and Kinahan, P. (2009). Clinical imaging characteristics of the positron emission mammography camera: Pem flex solo ii. *Journal of Nuclear Medicine*, **50**(10), 1666–1675.

Melcher, C. L. (2000). Scintillation crystals for pet. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine*, **41**(6), 1051.

Moliner, L., González, A., Soriano, A., Sánchez, F., Correcher, C., Orero, A., Carles, M., Vidal, L., Barberá, J., Caballero, L., *et al.* (2012). Design and evaluation of the mammi dedicated breast pet. *Medical Physics*, **39**, 5393.

Morris, E. A. (2002). Breast cancer imaging with mri. *Radiologic Clinics of North America*, **40**(3), 443–466.

Mosset, J.-B., Devroede, O., Krieguer, M., Rey, M., Vieira, J.-M., Jung, J., Kuntner, C., Streun, M., Ziemons, K., Auffray, E., *et al.* (2006). Development of an optimized lso/luyap phoswich detector head for the lausanne clearpet demonstrator. *Nuclear Science, IEEE Transactions on*, **53**(1), 25–29.

Moy, L., Ponzo, F., Noz, M. E., Maguire, G. Q., Murphy-Walcott, A. D., Deans, A. E., Kitazono, M. T., Travascio, L., and Kramer, E. L. (2007). Improving specificity of breast mri using prone pet and fused mri and pet 3d volume datasets. *Journal of Nuclear Medicine*, **48**(4), 528–537.

Nelson, W. R., Hirayama, H., and Rogers, D. W. (1985). Egs4 code system. Technical report, Stanford Linear Accelerator Center, Menlo Park, CA (USA).

Nvidia, C. (2011). Nvidia cuda programming guide.

Ortuño, J., Kontaxakis, G., Rubio, J., Guerra, P., and Santos, A. (2010). Efficient methodologies for system matrix modelling in iterative image reconstruction for rotating high-resolution pet. *Physics in medicine and biology*, **55**(7), 1833.

Panin, V. Y., Kehren, F., Michel, C., and Casey, M. (2006). Fully 3-d pet reconstruction with system matrix derived from point source measurements. *Medical Imaging, IEEE Transactions on*, **25**(7), 907–921.

Peng, H. and Levin, C. S. (2010). Design study of a high-resolution breast-dedicated pet system built from cadmium zinc telluride detectors. *Physics in medicine and biology*, **55**(9), 2761.

Peng, H. and S Levin, C. (2010). Recent developments in pet instrumentation. *Current Pharmaceutical Biotechnology*, **11**(6), 555–571.

Phelps, M. E., Hoffman, E. J., Huang, S.-C., and Ter-Pogossian, M. M. (1975). Effect of positron range on spatial resolution. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine*, **16**(7), 649.

Pratx, G. and Levin, C. (2011). Online detector response calculations for high-resolution pet image reconstruction. *Physics in medicine and biology*, **56**(13), 4023.

Pratx, G., Chinn, G., Olcott, P. D., and Levin, C. S. (2009). Fast, accurate and shift-varying line projections for iterative reconstruction using the gpu. *Medical Imaging, IEEE Transactions on*, **28**(3), 435–445.

Pratx, G., Surti, S., and Levin, C. (2011). Fast list-mode reconstruction for time-of-flight pet using graphics hardware. *Nuclear Science, IEEE Transactions on*, **58**(1), 105–109.

Qi, J., Leahy, R. M., Cherry, S. R., Chatziioannou, A., and Farquhar, T. H. (1998). High-resolution 3d bayesian image reconstruction using the micropet small-animal scanner. *Physics in medicine and biology*, **43**(4), 1001.

Rafecas, M., Boning, G., Pichler, B., Lorenz, E., Schwaiger, M., and Ziegler, S. (2001). A monte carlo study of high-resolution pet with granulated dual-layer detectors. *Nuclear Science, IEEE Transactions on*, **48**(4), 1490–1495.

Rafecas, M., Mosler, B., Dietz, M., Pogl, M., Stamatakis, A., McElroy, D. P., and Ziegler, S. I. (2004). Use of a monte carlo-based probability matrix for 3-d iterative reconstruction of madpet-ii data. *Nuclear Science, IEEE Transactions on*, **51**(5), 2597–2605.

Rahmim, A., Tang, J., Lodge, M., Lashkari, S., Ay, M. R., Lautamäki, R., Tsui, B., and Bengel, F. (2008). Analytic system matrix resolution modeling in pet: an application to rb-82 cardiac imaging. *Physics in medicine and biology*, **53**(21), 5947.

Rahmim, A., Qi, J., and Sossi, V. (2013). Resolution modeling in pet imaging: Theory, practice, benefits, and pitfalls. *Medical physics*, **40**, 064301.

Ramos, C. D., Erdi, Y. E., Gonen, M., Riedel, E., Yeung, H. W., Macapinlac, H. A., Chisin, R., and Larson, S. M. (2001). Fdg-pet standardized uptake values in normal anatomical structures using iterative reconstruction segmented attenuation correction and filtered back-projection. *European journal of nuclear medicine*, **28**(2), 155–164.

Ravindranath, B., Junnarkar, S., Purschke, M., Maramraju, S., Hong, X., Tomasi, D., Bennett, D., Cheng, K., Southekal, S., Stoll, S., *et al.* (2009). Results from prototype ii of the bnl simultaneous

pet-mri dedicated breast scanner. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pages 3315–3317. IEEE.

Raylman, R. R., Majewski, S., Smith, M. F., Proffitt, J., Hammond, W., Srinivasan, A., McKisson, J., Popov, V., Weisenberger, A., Judy, C. O., *et al.* (2008). The positron emission mammography/tomography breast imaging and biopsy system (pem/pet): design, construction and phantom-based measurements. *Physics in medicine and biology*, **53**(3), 637.

Rogers, D. (2006). Fifty years of monte carlo simulations for medical physics. *Physics in medicine and biology*, **51**(13), R287.

Rose, A. (1973). *Vision: human and electronic*. Plenum Press New York.

Sauter, A. W., Wehrl, H. F., Kolb, A., Judenhofer, M. S., and Pichler, B. J. (2010). Combined pet/mri: one step further in multimodality imaging. *Trends in molecular medicine*, **16**(11), 508–515.

Schaart, D. R., Seifert, S., van Dam, H. T., de Boer, M. R., Vinke, R., Dendooven, P., Lohner, H., and Beekman, F. J. (2008). First experiments with labr¡ inf¿ 3¡/inf¿: Ce crystals coupled directly to silicon photomultipliers for pet applications. In *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*, pages 3991–3994. IEEE.

Schmitt, D., Karuta, B., Carrier, C., and Lecomte, R. (1988). Fast point spread function computation from aperture functions in high-resolution positron emission tomography. *Medical Imaging, IEEE Transactions on*, **7**(1), 2–12.

Segars, W. P. and Tsui, B. (2009). Mcat to xcat: The evolution of 4-d computerized phantoms for imaging research. *Proceedings of the IEEE*, **97**(12), 1954–1968.

Shao, Y., Cherry, S., Farahani, K., Slates, R., Silverman, R., Meadors, K., Bowery, A., Siegel, S., Marsden, P., and Garlick, P. (1997). Development of a pet detector system compatible with mri/nmr systems. *Nuclear Science, IEEE Transactions on*, **44**(3), 1167–1171.

Shepp, L. A. and Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *Medical Imaging, IEEE Transactions on*, **1**(2), 113–122.

Siddon, R. L. (1985). Fast calculation of the exact radiological path for a three-dimensional ct array. *Medical physics*, **12**, 252.

Sohn, C. (1992). Ultrasound diagnosis of blood flow in benign and malignant breast tumors. *Geburtshilfe und Frauenheilkunde*, **52**, 397.

Sorenson, J. A., Phelps, M. E., and Cherry, S. R. (1987). *Physics in nuclear medicine*. Grune & Stratton Orlando.

Soret, M., Bacharach, S. L., and Buvat, I. (2007). Partial-volume effect in pet tumor imaging. *Journal of Nuclear Medicine*, **48**(6), 932–945.

Spanoudaki, V. C. and Levin, C. S. (2010). Photo-detectors for time of flight positron emission tomography (tof-pet). *Sensors*, **10**(11), 10484–10505.

Strother, S., Casey, M., and Hoffman, E. (1990). Measuring pet scanner sensitivity: relating countrates to image signal-to-noise ratios using noise equivalents counts. *Nuclear Science, IEEE Transactions on*, **37**(2), 783–788.

Strul, D., Slates, R., Dahlbom, M., Cherry, S. R., and Marsden, P. (2003). An improved analytical detector response function model for multilayer small-diameter pet scanners. *Physics in medicine and biology*, **48**(8), 979.

Stute, S., Benoit, D., Martineau, A., Rehfeld, N., and Buvat, I. (2011). A method for accurate modelling of the crystal response function at a crystal sub-level applied to pet reconstruction. *Physics in medicine and biology*, **56**(3), 793.

Sureau, F. C., Reader, A. J., Comtat, C., Leroy, C., Ribeiro, M.-J., Buvat, I., and Trébossen, R. (2008). Impact of image-space resolution modeling for studies with the high-resolution research tomograph. *Journal of Nuclear Medicine*, **49**(6), 1000–1008.

Surti, S. and Karp, J. (2008). Design considerations for a limited angle, dedicated breast, tof pet scanner. *Physics in Medicine and Biology*, **53**(11), 2911.

Surti, S. and Karp, J. S. (2007). Design considerations for a limited-angle, dedicated breast, tof pet scanner. In *Nuclear Science Symposium Conference Record, 2007. NSS'07. IEEE*, volume 5, pages 3667–3672. IEEE.

Thompson, C., Moreno-Cantu, J., and Picard, Y. (1992). Petsim: Monte carlo simulation of all sensitivity and resolution parameters of cylindrical positron imaging systems. *Physics in medicine and biology*, **37**(3), 731.

Tong, S., Alessio, A., and Kinahan, P. (2010). Noise and signal properties in psf-based fully 3d pet image reconstruction: an experimental evaluation. *Physics in medicine and biology*, **55**(5), 1453.

Von Schulthess, G. *et al.* (2000). Cost considerations regarding an integrated ct-pet system. *European radiology*, **10**, S377–S380.

Wang, G.-C., Huber, J., Moses, W., Qi, J., and Choong, W.-S. (2006). Characterization of the lbnl pem camera. *Nuclear Science, IEEE Transactions on*, **53**(3), 1129–1135.

Warner, E., Plewes, D., Shumak, R., Catzavelos, G., Di Prospero, L., Yaffe, M., Goel, V., Ramsay, E., Chart, P., Cole, D., *et al.* (2001). Comparison of breast magnetic resonance imaging, mammography, and ultrasound for surveillance of women at high risk for hereditary breast cancer. *Journal of Clinical Oncology*, **19**(15), 3524–3531.

Werner, M. K., Schmidt, H., and Schwenzer, N. F. (2012). Mr/pet: a new challenge in hybrid imaging. *American Journal of Roentgenology*, **199**(2), 272–277.

Wiant, D., Gersh, J., Bennett, M., and Bourland, J. (2010). Evaluation of the spatial dependence of the point spread function in 2d pet image reconstruction using lor-osem. *Medical physics*, **37**, 1169.

Wu, D. and Gambhir, S. S. (2003). Positron emission tomography in diagnosis and management of invasive breast cancer: current status and future perspectives. *Clinical breast cancer*, **4**, S55–S63.

Wu, Y., Bowen, S. L., Yang, K., Packard, N., Fu, L., Burkett Jr, G., Qi, J., Boone, J. M., Cherry, S. R., and Badawi, R. D. (2009). Pet characteristics of a dedicated breast pet/ct scanner prototype. *Physics in medicine and biology*, **54**(13), 4273.

Yamaya, T., Hagiwara, N., Obi, T., Yamaguchi, M., Ohyama, N., Kitamura, K., Hasegawa, T., Haneishi, H., Yoshida, E., Inadama, N., *et al.* (2005). Transaxial system models for jpet-d4 image reconstruction. *Physics in medicine and biology*, **50**(22), 5339.

Yang, X., Downie, E., Farrell, T., and Peng, H. (2013). Study of light transport inside scintillation crystals for pet detectors. *Physics in medicine and biology*, **58**(7), 2143.

Zaidi, H. and Hasegawa, B. (2003). Determination of the attenuation map in emission tomography. *Journal of Nuclear Medicine*, **44**(2), 291–315.

Zhang, J., Olcott, P. D., Chinn, G., Foudray, A. M., and Levin, C. S. (2007). Study of the performance of a novel 1 mm resolution dual-panel pet camera design dedicated to breast cancer imaging using monte carlo simulation. *Medical physics*, **34**, 689.

Zhang, L., Staelens, S., Van Holen, R., De Beenhouwer, J., Verhaeghe, J., Kawrakow, I., and Vandenberghe, S. (2010). Fast and memory-efficient monte carlo-based image reconstruction for whole-body pet. *Medical physics*, **37**, 3667.