Compton Imaging Algorithms for Position-Sensitive Gamma-Ray Detectors in the Presence of Motion

by

Jason M. Jaworski

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Nuclear Engineering and Radiological Sciences)
in The University of Michigan
2013

Doctoral Committee:

Professor Zhong He, Chair
Professor Jeffrey A. Fessler
Emeritus Professor Glenn F. Knoll
Associate Professor Sara A. Pozzi
Scott J. Wilderman

Where it all began

To my family

ACKNOWLEDGEMENTS

First I would like to thank my advisor, Professor Zhong He, for his guidance and support over the last five years. Without him, this work would not have been possible. I would also like to thank Professor Jeff Fessler for his invaluable guidance and discussions on model-based image reconstructions. I am grateful for the comments and suggestions of Professor Glenn Knoll during group meetings throughout my tenure at Michigan, and I appreciate the suggestions of Professor Sara Pozzi and Dr. Scott Wilderman who also agreed to be part of my committee. I am humbled to have such a distinguished group of experts serve on my dissertation committee.

I owe a great debt to the entire Orion group, past and present, for their excellent work and suggestions that made my research possible. The many discussions with members of the imaging group including Dr. Weiyi Wang, Dr. Christopher Wahl, Dr. Dan Lingenfelter, Sonal Kaye, Crystal Thrall, and Steven Brown were invaluable for my understanding and education. Of course, the imaging group would be a bunch of theorists without the core detector technology. Thanks to Dr. Feng Zhang, James Berry, Dr. Willy Kaye, and Andy Boucher for their critical effort on the detector hardware used for this work, and thanks to all of the other Orion members. I feel truly blessed to have been part of such a brilliant group of researchers.

Thanks to my many teachers and mentors whose encouragement and guidance led me to the nuclear engineering field and the Orion group. My interest in mathematics and science only grew as a result of my exceptional junior and senior high school teachers including Mrs. Fenn, Mrs. Becker, and especially Mrs. Jones in whose

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

CHAPTER I

Introduction

Radiation is all around us. Most people are surprised to discover that the bananas they eat are (slightly) radioactive or that the concrete walls in our buildings and the roads and sidewalks under our feet also emit radiation. Natural-background radiation is present in the Earth from uranium and its daughter products, and it is present in the air from Radon. Cosmic rays are constantly showering us from outer space. However, all of these sources of radiation are relatively weak, and most people never realize or care that they are constantly absorbing these background emissions. Radiation is not something to be feared. When understood and properly handled, radiation can be an effective tool that can power a city, sanitize food, kill cancerous cells, or detect and locate radiological threats.

1.1 Gamma-Ray Radiation

This work focuses on one specific type of radiation known as gamma rays which are part of the electromagnetic spectrum. As seen in Figure 1.1, electromagnetic radiation is a blanket term for may types of radiation including radio waves, microwaves, visible light, X rays, and gamma rays. The main distinction between these parts of the spectrum is the energy of the radiation. For example, radio waves have very low energy per photon (about 10^{-9} to 10^{-6} eV); whereas gamma rays have very high
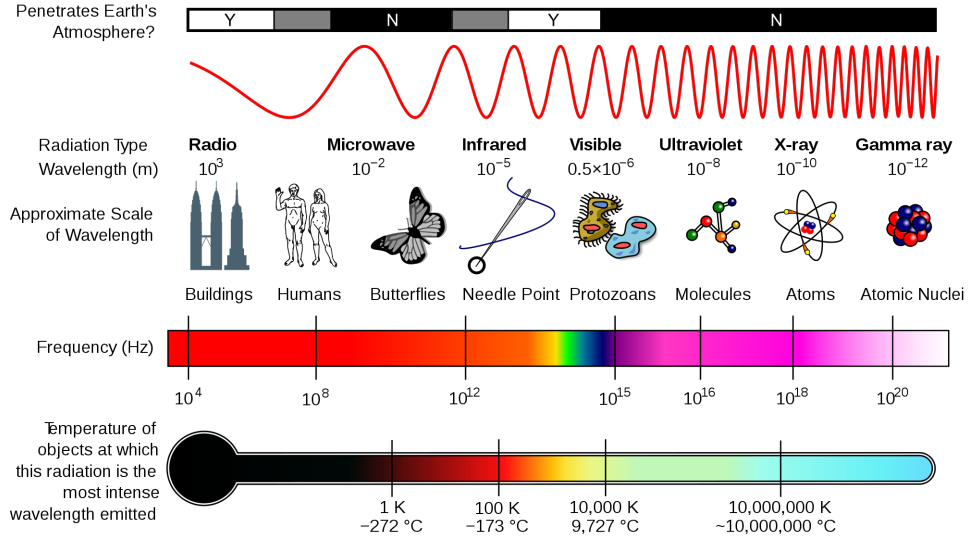
Figure 1.1: Electromagnetic spectrum [1].

energy (about 10^6 eV and up). The energy of the photons (individual particles of electromagnetic radiation) also determines how those photons interact with matter.

There are three main gamma-ray interactions with matter that are important for this work: photoelectric absorption, Compton scattering, and pair production. Figure 1.2 shows the relative probability of each type of interaction to occur as a function of energy in CdZnTe. At low energies the photoelectric effect dominates, and as the photon energy increases, Compton scattering is more likely. At energies much higher than 1.022 MeV, pair production is the dominant interaction type. Although the plot shows the probability for CdZnTe, which is the material of the detectors used in this work, the trends are similar for all materials.

The photoelectric effect occurs when a photon interacts with and excites an atom enough to eject an electron from the orbit of that atom. The photon deposits all of its energy then disappears, and the excess energy from the photon after the excitation of the atom is converted into kinetic energy of the ejected electron.

Compton scattering occurs when the gamma ray interacts with an electron directly, whether it is bound to an atom or not. Here the electron absorbs the full
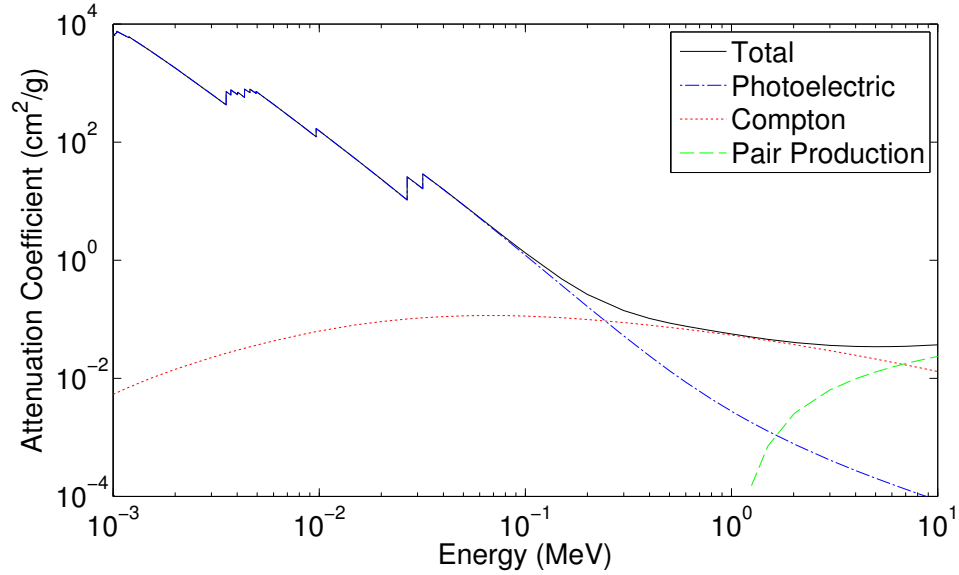
Figure 1.2: Gamma-ray attenuation coefficient (interaction probability) for CdZnTe [2].

energy of the photon initially. However, unlike photoelectric effect, a second photon of lower energy is emitted in a different direction than the original photon. The electron then carries the difference in energy between the scattered photon and the incident photon. The physics behind this interaction mechanism are described in more detail in Chapter III.

Pair production is only possible when the incident photon has an energy greater than 1.022 MeV (2x the rest mass of an electron). When a photon of adequate energy interacts with the electromagnetic field of the nucleus of an atom, the photon is converted into an electron and a positron (an anti-electron). The energy of the original photon exceeding 1.022 MeV is shared between the electron and positron. After the electron bounces around and loses its energy, it simply is absorbed by the surrounding material and is bound by an atom. However, when the positron slows down, it will annihilate with an electron and emit two new 511 keV photons.

1.2 Gamma-Ray Detection and Applications

Despite the fact that these gamma-ray interactions are constantly occurring all around us and even inside of us, humans cannot sense gamma-rays (or most other forms of radiation) without the help of specialized detection equipment. Therefore, many types of detectors have been developed over the years that allow us to measure this radiation [3,4]. Most detectors work by either directly or indirectly converting the gamma-ray radiation into electric charge. The charge is then measured and analyzed to determine the amount or intensity of the gamma rays, their energy, and even the location (or distribution) of their source.

Determining the location or distribution of gamma-ray sources is called gamma-ray imaging. Many applications including nuclear medicine, decontamination, border inspection, and nuclear nonproliferation can benefit from the use of gamma-ray imaging. In each of these applications, one similarity is the potential for moving sources and/or moving detectors. During medical procedures patients (or at least their organs) are moving. In decontamination procedures it would be convenient to move through a space with a detector to determine what objects show signs of contamination. Vehicles which could potentially carry illicit nuclear material are in motion as they travel to and from the border, and nuclear material travels through production facilities around the world.

Compensating for unknown source motion can be somewhat of a nuisance when performing image reconstruction, but a detector in motion has the potential to reconstruct 3-D tomographic radiation images. Tomographic imaging gives the user more information about the extent of the source and allows one to pinpoint the true 3-D position of the source instead of just the direction of the source. This work discusses the reconstruction of 3-D images but focuses on accounting for motion and using it to perform better image reconstruction.

1.3 Overview

Chapter II briefly describes the specific type of gamma-ray detectors used for this work, pixelated CdZnTe. Pixelated CdZnTe can measure both the energies of gamma-ray interactions and the positions of those interactions. It is the combination of energy and position information that enables one to perform the gamma-ray imaging described in this work.

Chapter III introduces the reader to Compton imaging and the reconstruction methods used in this work. Many before me have written similar summaries on Compton image reconstruction focusing on mathematical models, uncertainty calculations, and applications [5–7]. The focus of this review is more pragmatic. Important assumptions and options for the reconstruction methods and their effects on the resulting images follow a brief discussion on the basics of Compton imaging.

Chapter IV expands upon Chapter III to include 3-D image reconstructions. Several different methods of producing 3-D images are discussed. The final method, which uses moving detectors, introduces the first image reconstruction used in this work that involves motion.

Chapter V describes a novel reconstruction method for imaging sources with known motion. After a basic description of the benefits of the motion compensation scheme, this chapter focuses on the mathematical model of the reconstruction methods followed by imaging results reconstructed by this new method using experimental data.

The analysis in Chapter VI drops the requirement of knowing the path of the source motion. In crowded areas with complicated motion, reconstructing Compton images as a function of time becomes simpler than compensating for the motion of many moving objects. However, this time-dependent reconstruction introduces its own challenges which are addressed in this chapter.

Finally, Chapter VII discusses some areas of the image reconstructions that need

more work or other directions in which this type of work may make an impact.

1.4   Contributions of this Work

The main novel contributions of this work are contained in Chapters IV-VI. The focus of these contributions is to account for the presence of motion in the imaging algorithms previously developed for static scenarios. The 3-D image-reconstruction algorithms we describe in Chapter IV allow one to create tomographic images with a moving detector system. In this case, the motion of the detector generates the required parallax to reconstruct the 3-D distribution of radiation sources.

When instead the sources are in motion and the detector is stationary, this work focuses on two different novel approaches to account for this motion which would otherwise cause blur in the reconstructed image. The first reconstruction method is for situations where the source motion is known or can be estimated, *e.g.,* by using optical tracking methods for cars at a border crossing. The addition of new spatial-distribution bins that track the source motion allow the algorithm to faithfully reconstruct the source distribution for multiple moving and stationary sources.

For the situation where the source motion is unknown, we add the time domain to the image reconstruction and take several different approaches to reduce the statistical noise expected for low count rate scenarios. The first approach is to add a regularization parameter to encourage the reconstructed image to be smooth in the time domain. The other techniques involve using moving-average and median filtering to reduce statistical fluctuation in the time domain which ultimately lead to a reduction in artifacts found in the imaging domain.

All of the image reconstruction algorithms described in this work are implemented in C++ and have been included in the UMImaging code package [8].

CHAPTER II

Pixelated CdZnTe Detectors

Although the focus of this work is Compton image reconstruction, it relies on the data recorded by gamma-ray detectors. In an appropriate configuration, many types of gamma-ray detectors can produce the data required to perform Compton imaging. However, pixelated CdZnTe detectors have several advantages over other types of detectors including room-temperature operation, excellent energy resolution, and 3-D interaction position sensitivity. The following sections briefly describe the detector operation and configuration and are followed by a description of a Geant4 model of the detector system.

2.1 CdZnTe Operation

CdZnTe is a compound-semiconductor detector with a relatively wide band-gap of 1.6 eV [3, pg. 492]. Like other semiconductor detectors, a bias voltage is applied to the detector volume during operation to create a depletion region extending through most of the volume in which all of the free electric carriers (electrons and holes) have been swept out of the detector. The large bandgap means that at room temperature the probability of thermal excitation of free charge carriers in this depletion region is much lower than detector materials like high purity germanium (HPGe). This low thermal carrier generation and high resistivity allow the detectors to operate at room

temperature without excess noise from thermal carrier generation and the leakage current through the bulk. This room temperature operation is one key advantage CdZnTe has over materials like HPGe.

As discussed in the previous chapter, when gamma-rays interact with matter they excite and ionize atoms. In semiconductors, this ionization process generates a quantity of electron-hole pairs proportional to the energy deposited by the gamma ray. Normally these would simply recombine in the absence of an electric field, but the bias voltage applied to the detector causes an electric field to form which attracts the electrons to the anode surface and holes to the cathode surface. The goal is to measure the amount of charge generated by the interaction to determine the energy of the incident photon and the position at which the interaction occurred by observing the electrical signals produced by the charge moving through the detector.

2.2 Charge Collection and Measurement

A key difference between CdZnTe and materials like HPGe is poor hole mobility. HPGe detectors have relatively high mobility for electrons and holes allowing them to use the induced signal from both carriers moving through the bulk. CdZnTe has poor hole mobility compared to its electron mobility to the point that the holes are essentially stationary during the time in which the electrons are completely collected by the anode. This lack of hole motion means that CdZnTe detectors must rely solely on the motion of electrons to determine the deposited energy.

Before we can analyze how the lack of hole motion affects the measured signal, we must have an understanding of the charge-induction process. The Shockely-Ramo Theorem [9–11] states that the charge Q induced on an electrode by point charge q is

$$Q = -q\big[\phi_0(\vec{r}_2) - \phi_0(\vec{r}_1)\big] \tag{2.1}$$

where ϕ_0 is the weighting potential, $\vec{r_1}$ is the location where the charge is created, and $\vec{r_2}$ is the final location of the charge. The weighting potential is unique to each electrode and is calculated by setting the electrode of interest to unit potential while grounding all other electrodes and solving Poisson's equation ignoring space charge (including q)

$$\nabla \phi_0 = -\frac{\rho}{\epsilon} = 0 \tag{2.2}$$

where ρ is the space charge and ϵ is the permittivity of the medium. If there is a magnetic field present, this analysis assumes that it is constant. The implication of this result is powerful. It means that the charge induced on an electrode depends on the change of the weighting potential between the time the charge is generated and when it is collected but is completely independent of the actual applied bias, the path taken by the charge, and any space charge that may be present.

2.2.1 Planar Detectors

Consider a cubic detector geometry with a simple planar cathode and anode seen in Figure 2.1. The weighting potential calculated using (2.2) for both electrodes is shown as a function of depth in Figure 2.2. For simple planar electrodes, the weighting potential is uniform in the lateral directions. Assume that the detector material is HPGe and a gamma-ray interacts at some depth d in the detector generating n electron-hole pairs which have unit charge $\pm q_0$. If we focus on the induced charge on the cathode, (2.1) combined with Figure 2.2 predicts that the induced charge from the holes is

$$Q_h = (1 - d/D)nq_0 \tag{2.3}$$

where D is the total depth of the detector. Similarly, the induced charge from the electrons is

$$Q_e = (0 - d/D)n(-q_0) \tag{2.4}$$

9
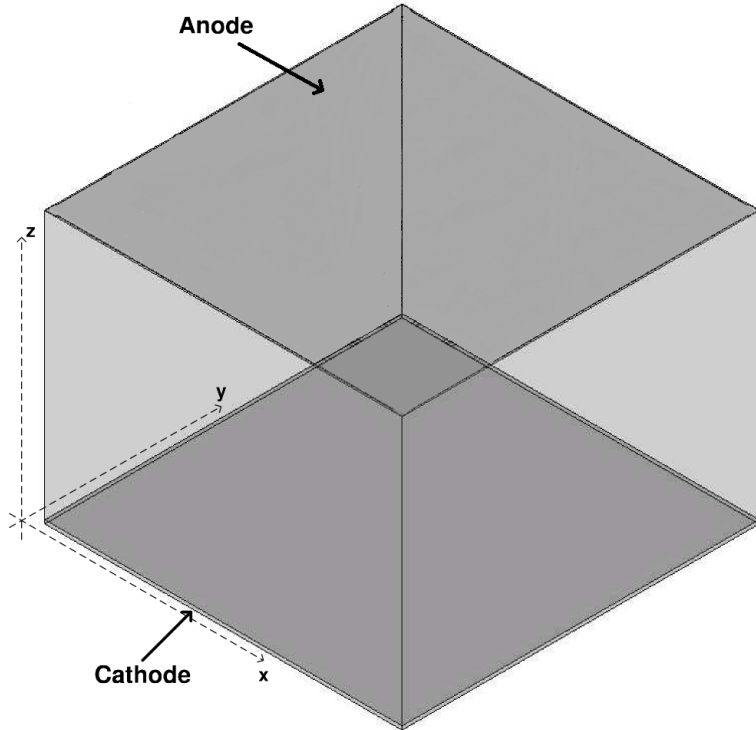
Figure 2.1: Diagram of a 2×2×1.5 cm^3 planar detector.

since electrons have a negative charge, and the total induced charge is

$$Q_t = Q_h + Q_e = nq_0 \qquad (2.5)$$

which is proportional to the total energy deposited by the gamma-ray interaction. However, if we instead assume that the material is CdZnTe, there is no hole movement and the total induced charge is only $(d/D)nq_0$. Now the signal is proportional to both the deposited energy and the depth-of-interaction (DOI).

2.2.2 Pixelated Detectors

Several different methods of overcoming this depth dependency (all of which take advantage of the Shockley-Ramo theorem) have been explored [12–14], but this work uses a pixelated anode approach [15]. Figure 2.3 shows a diagram of a pixelated CdZnTe detector with dimensions of $2 \times 2 \times 1.5$ cm^3. The cathode is still a simple
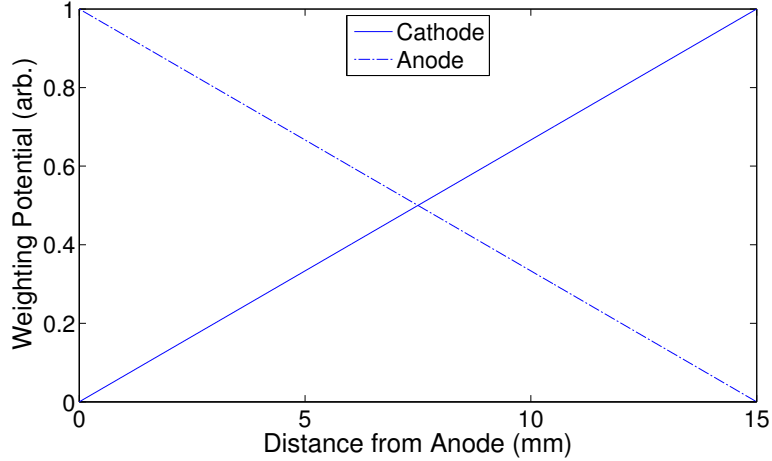
Figure 2.2: Weighting potential as a function of depth for the anode and cathode of a planar detector.

planar electrode as in the previous example, but the anode now consists of 11×11 square pixels with a pixel pitch of 1.72 mm and a grid electrode that surrounds each of the pixels. With an appropriate negative bias, the grid helps steer the charge to the pixels for better charge collection.

Figure 2.4 shows the weighting potential as a function of depth underneath the pixel that collects the charge for the cathode, the collecting anode pixel, and a nearby non-collecting anode pixel. The weighting potential for the cathode is exactly the same as it was for the planar detector. However, because of the small-pixel effect [16], the weighting potential for the anode pixels stays almost flat for most of the bulk of the detector. At about a distance of one pixel pitch from the anode side, the weighting potential for the collecting anode increases sharply to one, but the weighting potential for the non-collecting anode initially rises and then drops to zero.

The anode pixels are relatively insensitive to the movement of charge in most of the detector volume since the change in their weighting potential is almost negligible until the sharp rise near the anode surface. Thus, assuming electron trapping is negligible, the induced signal on the collecting pixel is almost independent of the DOI unless the interaction occurs within a pixel pitch of the anode surface. In this
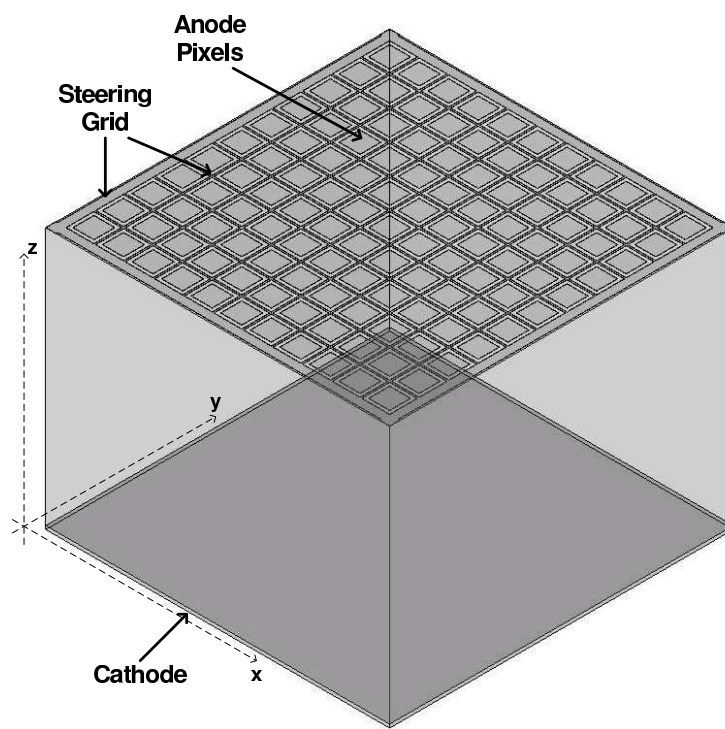
Figure 2.3: Diagram of a 2×2×1.5 cm^3 pixelated CdZnTe detector.
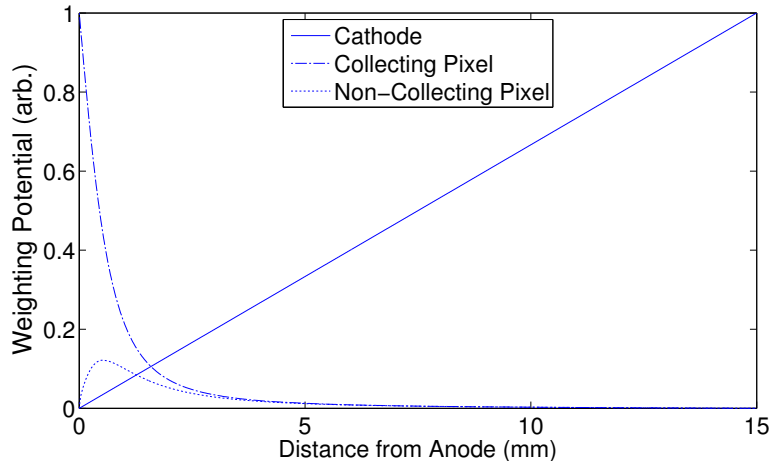
Figure 2.4: Weighting potential as a function of depth for the anode pixels and cathode of a pixelated detector corresponding to the lateral position directly underneath the collecting pixel.

simple example where all of the charge is collected by one pixel, the non-collecting pixels will result in a slightly negative signal depending on the DOI.

Now we have the amplitude of the cathode signal, which is proportional to the DOI and the energy deposited, and the amplitude of the collecting-anode signal, which is only proportional to the energy deposited. If we compute the ratio of the cathode signal amplitude to the collecting-anode signal amplitude, a method pioneered by He et al. [17], we get the DOI. With the depth information and the 2-D position of the collecting-anode pixel, we can calculate the 3-D position of the gamma-ray interaction in the detector volume. This method works when only one interaction occurs in the detector volume during the signal integration time. However, later work introduced the measurement of electron-drift time to determine the DOI of multiple interactions resulting from a single incident photon [18].

2.3   Non-Ideal Characteristics

Of course in reality there are many non-ideal characteristics about the system that cause reduced performance and ultimately affect the images reconstructed using the

13

methods described in the following chapters. Some of these limitations are a result of the detector itself, and others can be attributed to the readout electronics. The next two sections describe a few of these limitations.

2.3.1 Detector Limitations

Pixel jumping is an effect recently observed in pixelated CdZnTe detectors [19] that refers to the process of charge moving laterally in the bulk of the detector so that it is collected by a pixel other than the pixel under which it was originally generated. The obvious implication here is that the true position of the interaction differs from that recorded by the detector system. This inaccurate position information can cause problems for image-reconstruction methods that require this information. Fortunately, Compton imaging is less affected by this phenomenon than other imaging modalities, but it is still adds to the uncertainty of the image reconstruction.

Dead layers and/or pixels are present to some extent in all pixelated CdZnTe detectors. The volume just below the anode surface is essentially a dead region because of the relatively small change in the weighting potential between the initial and final positions of the electrons. The detector cannot sense gamma-ray interactions in this region, and therefore it does not contribute to the active volume of the detector. This dead layer is present for every pixelated detector and causes a loss of efficiency and some non-uniformity in the detector response.

Dead anode pixels are not uncommon and can vary in severity. Some pixels are dead to interactions of all energies, and some are just dead to low energy interactions (on the order of 60 keV). These dead pixels are usually a result of poor fabrication of the anode pixels and again cause non-uniformities and loss of efficiency.

Finally, some detectors have regions of active volume in which the electron drift velocity is non-uniform [20]. Since we use that drift time to reconstruct the DOI for events where more than one anode pixel collects charge, if the magnitude of the

14

electron-drift velocity is relatively low in a particular region, the uncertainty in the reconstructed DOI increases because of the slower initial rise of the induced signal. This degradation in depth resolution can cause significant degradation in the quality of the image reconstruction.

2.3.2 Readout Electronic Limitations

The VAS_UM/TAT4 application specific integrated circuit (ASIC) designed by Gamma-Medica Ideas [21] is at the front end of the readout electronics for the experimental results shown in this work. This piece of circuitry contains 129 channels, each of which reports the pulse amplitude and timing information which are then read out by an FPGA board and eventually to a computer for analysis. Despite the excellent performance this ASIC technology can achieve, there are some limitations that affect the imaging performance.

As with any circuit, some level of electronic noise and interference between signals is present in the output from the ASIC. This noise and interference ultimately manifests itself as uncertainties in the measurement of the deposited energies and interaction depths. The uncertainty in the energy measurement for the current system does not limit the current imaging reconstructions, but the uncertainty in the DOI does have some significant effect on the image quality for reasons discussed in Chapter III.

"Charge sharing" refers to when an electron cloud created by a gamma-ray interaction is collected by two or more pixels. This sharing can happen when an interaction occurs midway between two pixels or when the charge cloud is so large it spills over to neighboring pixels. When charge is shared between pixels, the uncertainty in the energy measurement is increased because of the addition of noise in multiple ASIC channels. There is also the possibility that the charged shared on a pixel is too small to trigger the channel which would cause that energy information

to be lost entirely. Charge sharing also makes it difficult to differentiate between true multiple-interaction events and single-interaction charge-sharing events. However, recently Wang et al. [22] have shown that the best way to handle these events is to simply combine the information from the triggered pixels by summing the energies and setting the interaction location to be the energy-weighted centroid of each individually-reconstructed position.

One other limitation of the VAS_UM/TAT4 ASIC is that it can only measure a single interaction under any one pixel. Because the induced-charge signal is sent through a shaping filter, when two interactions occur under a single anode pixel, the induced charge simply piles up and the resulting shaped signal is roughly proportional to the sum of the energies deposited. This effect reduces the number of events that can be used for imaging, but this loss of events is relatively small affecting about 5% of single-pixel events [23] (although this percentage does as a function of source energy and position).

2.4   18-Detector Array

To improve overall efficiency, we made an array consisting of 18 $2\times2\times1.5$-cm^3 pixelated CdZnTe detectors. The geometry of the detectors which consists of two 3×3 planes of detectors is shown in Figure 2.5. The ASICs are connected to the anode side of the detectors and are depicted as light grey squares on mounted on the green printed circuit boards. This arrangement places the dead layer near the anode surface on the inside of the array so that any low energy gamma-rays interacting near the outer surface of the array can be detected.

The system operates at room temperature and triggers as a single detector unit. The triggering scheme allows a photon that scatters between detectors or detector planes to be treated as a single multiple-pixel event instead of several individual events. This improves full-energy deposition efficiency and imaging efficiency since
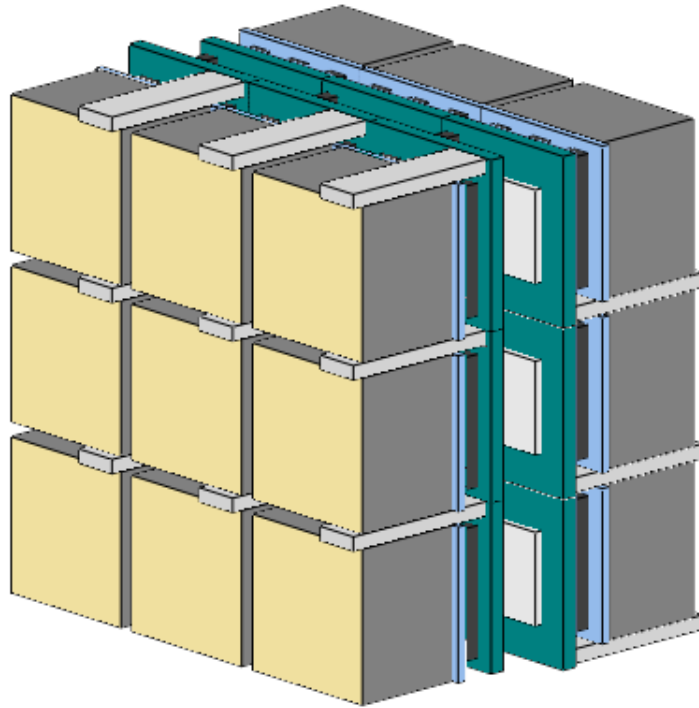
Figure 2.5: Diagram of the 18-detector array. The pixelated detectors are arranged in two 3×3 planes with the cathode surfaces facing out to increase low-energy gamma-ray detection efficiency.

Compton imaging requires the use of multiple-interaction events. It can achieve better than 1% energy resolution at 662 keV for single-pixel events and about 1.4% energy resolution at 662 keV for all events combined and has a relative efficiency of about 30% at 1333 keV compared to a 3×3 inch NaI(Tl) scintillation detector.

This detector array is the main source of experimental data for the results presented in this work.

2.5   Geant4 Simulation

Some analysis in this work requires the use of simulated data. A simulation package written in the C++ programing language and based on the Geant4 toolkit [24] is used to generate gamma-ray interaction-event data similar to that recorded by real pixelated CdZnTe detectors. Many geometries including the 18-detector array are supported as well as many source types including parallel beams and isotropic point sources that are stationary or in motion.

The simulation uses physics packages that include not only the gamma-ray interactions described in the previous chapter, but also the electron cloud production. The electric field and charge transport are not included in the model, but pixelation and charge sharing effects are included as well as realistic energy and DOI uncertainties.

CHAPTER III

Gamma-Ray Imaging

Gamma-ray imaging is fundamentally different than optical imaging because gamma rays cannot be focused easily. When someone takes a picture of the sunset with their camera, the light incident on the camera lens is bent and directed toward the film or light sensor. However, despite some recent work [25], it is still impractical to image gamma rays this way due to the nature of gamma-ray interactions with matter. Because of this key difference, gamma-ray imaging requires a different approach.

In general, gamma-ray detectors record information about the photon interactions inside the detector volume, but the data itself is not natively an image (except in transmission imaging like an x-ray of a broken leg). Therefore, the recorded data has to be processed in order to reconstruct a useful image. This chapter describes two Compton image reconstruction methods that are used to generate meaningful images.

3.1 Definitions

The goal of the following reconstruction methods is to use the list-mode data (event-by-event data that includes the energies and positions of each photon interaction) recorded by the detector, $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, to reconstruct the source intensity as a function of direction and energy, $\lambda(\vec{r}, E)$, where E is the incident photon energy and $\vec{r}$ is the direction from which the photon originated, and n is the number of recorded

events.

A typical event $\boldsymbol{v}_m$ recorded by the detector system contains the 3-D positions and energies of the individual photon interactions that define it. For example, a photon may enter the system, Compton scatter, and get absorbed. This event would contain the position and energy-deposition information at the scatter site as well as the position and energy deposition information at the absorption site.

In this chapter, the imaging domain is restricted to the 2-D directional space surrounding the detector. This space can be visualized as a sphere with the detector centered at the origin as seen in Figure 3.1(a). Spherical coordinates at a fixed radius are used to describe the space in terms of polar and azimuthal angle. The energy domain is typically a simple histogram of energies ranging anywhere from 0-3 MeV.

3.2 Simple Back Projection Reconstruction

The simple back projection (SBP) method uses the gamma-interaction data to project the possible incident directions (*e.g.,* a Compton cone) back into the imaging space. Generally, because of its simplicity (*i.e.,* it does not include calculations to estimate attenuation and interaction probabilities and it is not iterative) SBP is fast, but for Compton imaging the quality of the reconstructed source distribution is not as good as the more advanced reconstruction techniques.

3.2.1 Compton Imaging

Compton imaging is based on the kinematics of a photon that undergoes Compton scattering. Traditionally, Compton imaging systems required two 2-D-position-sensitive planes of detectors: one to scatter the incident photons and one to absorb those scattered photons [26, 27]. However, Compton imaging using 3-D-position-sensitive detectors differs from traditional Compton imaging or even other imaging modalities like CT, PET, SPECT, or coded-aperture imaging, because it does not re-
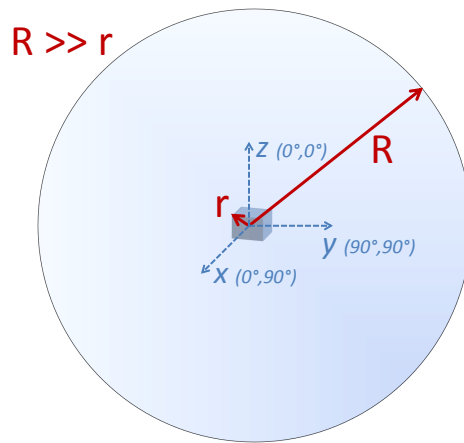
(a)
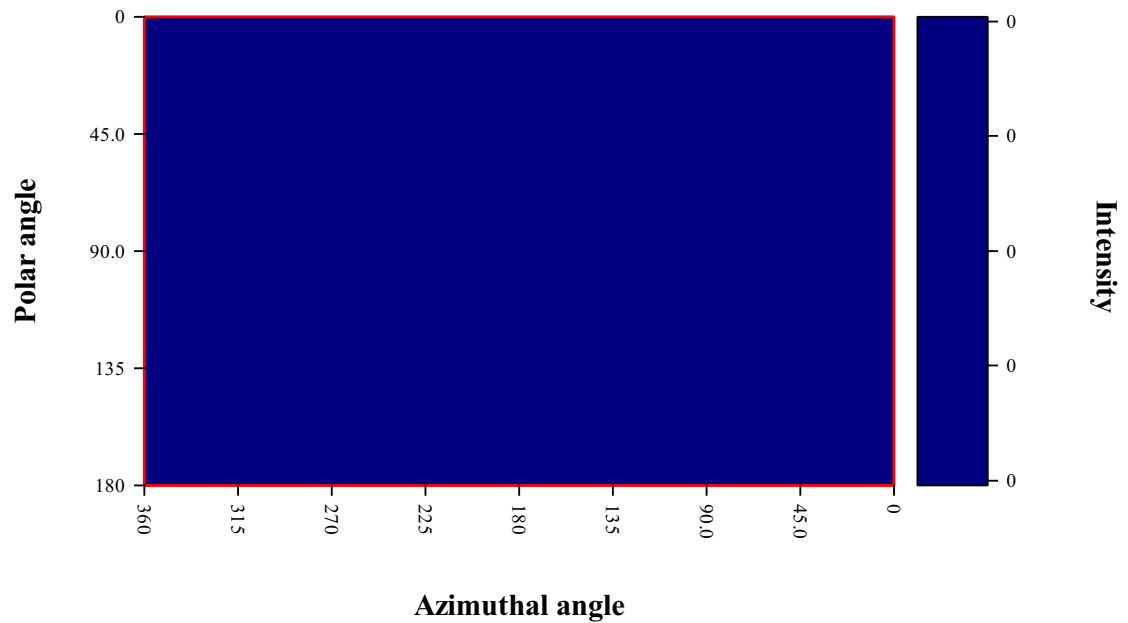
(b)

Figure 3.1: (a) Imaging domain depicted as a large sphere with the detector at the origin. (b) Imaging space represented as a Mercator projection in which the whole imaging sphere is represented using a standard rectangular mesh.

quire collimation or a specific detector geometry. A single pixelated CdZnTe detector (like those used in this work) can perform Compton imaging by itself because it can record the 3-D positions and energies of each photon interaction in the active volume. In fact, any gamma-ray spectrometer with the ability to record the 3-D locations of multiple-interaction events can be used as a Compton imager.

3.2.2 Kinematics

Assume that a gamma ray incident on the detector interacts through Compton scattering and that the scattered photon is then captured by a photoelectric absorption. A line drawn between the two interaction locations represents the axis of a cone whose surface defines the set of possible origins of the incident photon (see Figure 3.2). The half angle of the cone, θ, is determined by the kinematics of the Compton scatter interaction. Through conservation of energy and momentum and assuming that the photon interacts with an unbound electron at rest, one can derive the following simple expression:

$$E' = \frac{E_0}{1 + \frac{E_0}{m_0 c^2}(1 - \cos(\theta))} \tag{3.1}$$

where E_0 is the energy of the incident photon, E' is the energy of the scattered photon, m_0 is the rest mass of an electron, and c is the speed of light. In this example, the sum of the energies deposited in the two interactions is equal to the energy of the incident photon, E_0, and the energy deposited in the photoelectric absorption is the energy of the scattered photon, E'. One can then easily solve for the scatter angle θ.

As mentioned earlier, this calculation only determines the half angle of a cone that defines a set of possible source directions. A single event alone will not reveal the direction of a single point source. This ambiguity is the result of not knowing the initial direction of the scattered Compton electron, ϕ. If this direction is known, the Compton cone collapses to a line connecting the initial interaction location and the source location.
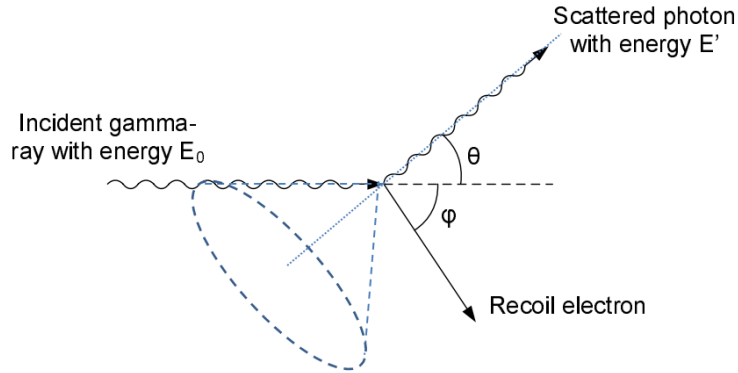
Figure 3.2: Example of a Compton scatter and Compton cone.

Unfortunately, tracking the initial direction of the scattered electron is difficult because of the sporadic path that electrons take. Until recently, the ability to record the scattered electron track was limited to gas detectors [28] where the low density reduces the probability of the electron to undergo a high angle scatter. Ironically, however, their low density also makes them poor gamma-ray detectors since gas has such a low stopping power. Some recent work [29] shows promise of recording the electron track in a solid state system using silicon charge-coupled devices (CCDs) which allow the reconstruction to restrict the possible set of incident directions to a small cone segment. However, this type of system requires an absorption plane like traditional Compton cameras, and the readout and processing of the CCD data is slow. Both of these factors reduce efficiency, but each event carries more information since much of the ambiguity of the Compton cone is eliminated.

For most Compton imaging detector systems, including the CdZnTe based detector systems used for this work, the electron track information is unknown. In this case, several Compton cones must be reconstructed to reveal the true direction of a point source. As seen in Figure 3.3, all of the reconstructed Compton cones intersect in the source direction, but some also overlap in other chance locations. Thus, because of these chance intersections combined with measurement uncertainties and interference from other sources and background, many cones are required to determine the true
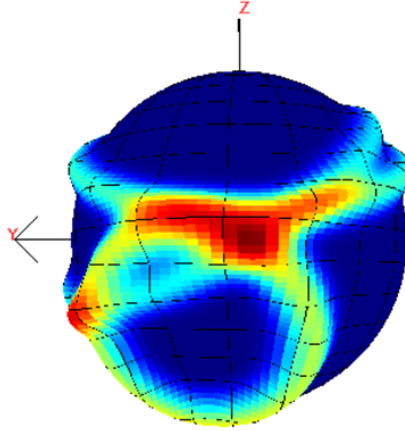
Figure 3.3: Compton cones overlap to reveal the source direction.

source direction in practice.

3.2.3 Full-Energy Assumption

An important assumption made for this reconstruction is that the incident photon deposits all of its energy in the active volume of the detector. E_0 is unknown and is estimated by summing the energies of all of the interactions in the event. This assumption is obviously far from reality since a large fraction of incident photons deposit some energy and escape the detector system entirely (especially for relatively high energy photons and/or small detector volumes).

Therefore, when using SBP, it is important to use energy windowing to only reconstruct full energy photopeak events. This filtering can be set before the reconstruction has started if an isotope of interest is known *a priori.* However, a quick look at the energy spectrum of the recorded events may reveal other sources of interest, and the window could be adjusted to accommodate them. Also, if the source energy is known (or is revealed after a look at the spectrum), the reconstruction can be performed using that known energy for E_0 and the known energy minus the energy deposition of the initial Compton scatter for E' in 3.1. This substitution will result in the correct cone angle calculation despite the possibility that the full energy was not deposited

assuming the photon did not undergo any scattering before reaching the detector.

3.2.4   Sequence Reconstruction

One seemingly simple yet important piece of information required for Compton imaging is the ability to know which interaction occurred first and which one was second. The first interaction defines the vertex of the reconstructed cone, and the first interaction combined with the second interaction defines the cone axis. Although they carry important energy information, the interactions beyond the first and second are irrelevant for defining the geometry of the Compton cone. Unfortunately, the detectors used for this work cannot resolve the time-of-flight information to determine which interaction occurred first. The interactions happen almost simultaneously since the average distance between interactions is on the order of a few millimeters which corresponds to about a 10 ps time difference.

Without the time-of-flight information, there are $N!$ different possible interaction sequences, where N is the number of interactions in the event. To correctly reconstruct the Compton cone a sequence reconstruction algorithm must be used to estimate the correct interaction sequence. We use three different sequence reconstruction algorithms for this purpose: simple comparison, deterministic, and mean-squared difference (MSD). The following sections give a brief introduction to these methods.

3.2.4.1   Simple Comparison Method

The simple comparison algorithm is designed to work with 2-interaction events only. In this case there are only two possible sequences. Under the assumption that the incident photon deposited its full energy in the two interactions, one can use (3.1) to determine if a sequence is physically possible. Based on the kinematics, the maximum energy that a photon can deposit in a Compton scatter, E_{max}, occurs when

the scattering angle is 180°:

$$E_{max} = \frac{E_0}{1 + \frac{m_0 c^2}{2E_0}}. \tag{3.2}$$

The first step in the simple comparison method is to check both energy depositions to see if one of them deposits more energy than E_{max}. If one does, then that interaction must have been the second interaction, and the true sequence is found.

If both energies are below E_{max}, then another approach must be taken. At this point Xu [30] proposes comparing the two deposited energies and picking the one with the higher energy to be the first interaction. This proposal was based on a simulation that suggested a larger fraction of photons deposit more energy in their first interaction for 15 mm x 15 mm x 10 mm CdZnTe detectors. This result is not immediately obvious, but if more energy is deposited in the first interaction, less energy is left for the scattered photon, making it more likely to be captured. Despite the fact that the detectors used in this work are larger and are part of an array, the simple comparison method still works well for the larger geometry.

3.2.4.2 Deterministic Method

The deterministic sequence reconstruction method, unlike simple comparison, works for events with an arbitrary number of interactions. Of course, there is a practical limit since the number of possible sequences is $N!$, which quickly gets out of hand for events with more than just a few interactions. Luckily, for the photon energies of interest in this work (0-3 MeV), the probability of having more than 5 interactions is usually negligible for CdZnTe detectors.

This algorithm calculates the probability of each interaction sequence and chooses the sequence that has the highest likelihood of occurring. The factors used in the calculation include attenuation and scatter cross-section probabilities for interaction

positions $\vec{r}_1 \ldots \vec{r}_n$ and energies $E_1 \ldots E_n$:

$$
\begin{aligned}
\mathsf{Pr}_{\text{seq}} = & \mathsf{Pr}[\text{Incident photon reaches } \vec{r}_1] \cdot \\
& \mathsf{Pr}[\text{Compton scatter deposits energy } E_1] \cdot \\
& \mathsf{Pr}[\text{Scattered photon travels } \vec{r}_1 \to \vec{r}_2] \cdot \\
& \mathsf{Pr}[\text{Compton scatter deposits energy } E_2] \cdot \\
& \vdots \\
& \mathsf{Pr}[\text{Scattered photon travels } \vec{r}_{n-1} \to \vec{r}_n] \cdot \\
& \mathsf{Pr}[\text{Photoelectric absorption deposits energy } E_n],
\end{aligned}
\tag{3.3}
$$

For 2-interaction events, the Compton edge test is performed first to determine if one of the sequences is invalid.

In practice, the deterministic method performs slightly better than the simple comparison method, but it is slightly more computationally expensive. Thus, if speed is not of the utmost importance, the deterministic method should be used for 2-interaction events.

3.2.4.3   Mean Squared Difference Method

The mean squared difference sequence reconstruction method only works for events with three or more interactions (with the same practical upper limit as the deterministic method). It requires three or more interactions because it compares the scatter angles determined energetically using (3.1) to those determined geometrically by

$$
\cos \theta_r = \frac{(\vec{r}_2 - \vec{r}_1) \cdot (\vec{r}_3 - \vec{r}_2)}{\|(\vec{r}_2 - \vec{r}_1)\| \|(\vec{r}_3 - \vec{r}_2)\|}.
\tag{3.4}
$$

Thus, with only two interactions, no geometric calculation is possible.

The figure of merit (FOM) for each interaction sequence is described by Xu et

al. [30] to be the square of the difference between the scatter angle determined by energy and geometry divided by the variance in that difference:

$$FOM_{\text{seq}} = \frac{(\cos\theta_e - \cos\theta_r)^2}{\sigma^2(\cos\theta_e - \cos\theta_r)} \qquad (3.5)$$

where θ_e is the scatter angle calculated using the energy information, θ_r is the scatter angle calculated geometrically, and $\sigma^2(\cos\theta_e - \cos\theta_r)$ is the variance in the difference described in [30]. For events with more than three interactions, this ratio can be calculated for multiple scatter locations. In that case the FOM_{seq} is simply the product of those ratios. The sequence with the lowest figure of merit is selected as the "correct" sequence.

As shown by Thrall et al. [31], the MSD method performs better than the deterministic method and is also less computationally expensive. Thus, for events with three or more interactions, MSD is the sequence reconstruction method of choice.

3.2.5   Key Reconstruction Options

Xu [5] describes in detail how Compton cones are reconstructed when considering measurement uncertainties from the detector, and Section 3.3.2.2 describes the system response function to calculate the distribution of a single ring. The resulting ring distributions of each recorded photopeak event are then added together to reconstruct the SBP image. However, here we focus on several options one can set that affect both how the cone is drawn and what type of events are used in the reconstruction. These options can be set to fit a specific need when using this reconstruction method.

3.2.5.1   Cone Width

Determining the width of the back projected ring is one such option that can be tailored to a given problem. The standard ring width is calculated to correspond

to the uncertainty of the measurement. The two uncertainties that affect the simple back projection reconstruction are the energy uncertainty and the interaction position uncertainty. The recorded energy primarily affects is the cone angle, θ. In contrast, the interaction position uncertainties mostly affect the direction of the cone axis.

To demonstrate the relative importance of each type of uncertainty, an SBP reconstruction using rings with a constant narrow width is used to reconstruct simulated data that isolates each type of uncertainty. Figure 3.4(a) is the resulting image from an SBP reconstruction using data from a Geant4 simulation that reported exact interaction energy and position information of 662 keV photons emitted from a point incident on the detector array described in section 2.4. Figure 3.4(b) displays the SBP image reconstructed using data from the same Geant4 simulation with an added energy blur consistent with the experimental energy resolution of the CdZnTe detectors. Finally, Figure 3.4(c) shows the reconstructed SBP image using data from the Geant4 simulation with an added pixellation uncertainty and 0.5 mm depth blur on the interaction locations (but still perfect energy). Doppler broadening effects due to the initial motion of the Compton electron are ignored for this study.

One can see from the figures that the position uncertainty is the dominant source of blur in the reconstructed image. The angular resolution only degrades from about 2.5° to 6° FWHM when energy uncertainty is included, but the position uncertainty increases the resolution to about 30° FWHM. This result is expected since the average separation distance of interactions in CdZnTe from 662 keV incident photons is only a few millimeters. Remember that the pixel pitch of the detectors is 1.72 mm, and the depth uncertainty is about 0.5 mm. Thus, the position uncertainty is of the same order of magnitude as the separation distance of the interactions, yielding large uncertainty in the cone axis.

The inclusion of this uncertainty in the ring reconstruction is an essential part of the model-based reconstruction methods described later in this chapter; however, it
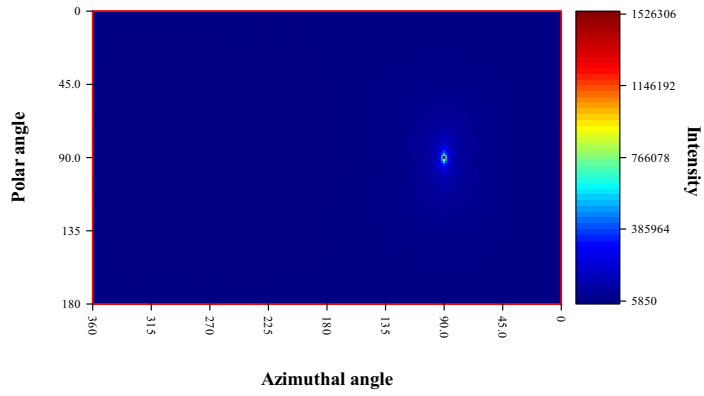
(a)
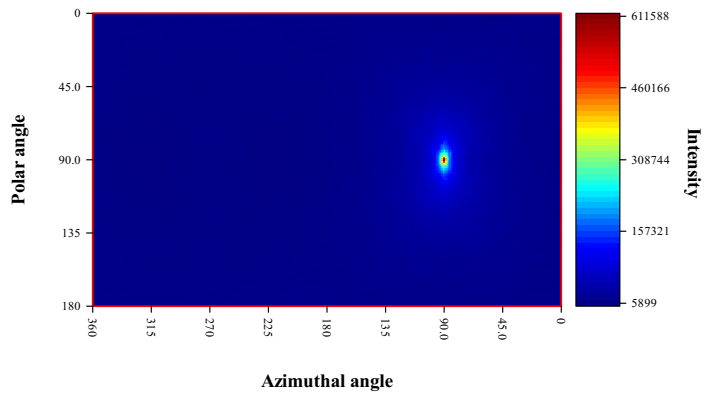
(b)
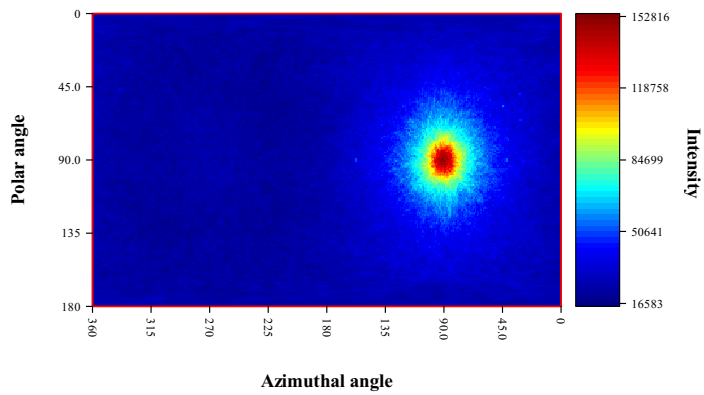
(c)

Figure 3.4: (a) SBP reconstruction with perfect energy and position information. (b) SBP reconstruction including energy uncertainty only. (c) SBP reconstruction including position uncertainty only.

may seem strange, intuitively, to add blur to the image by increasing the width of the rings. This added blur is analogous to adding a Gaussian distribution centered on the recorded energy bin in the pulse height spectrum instead of simply adding a count to the appropriate energy bin. In fact, the added blur actually degrades the angular resolution which is already poor.

The reason extra blur is added to the cones in the SBP reconstruction is that it allows the source direction to be more obvious with a limited number of counts. Figure 3.5 illustrates this effect. A similar effect may be accomplished by using a coarser imaging mesh, which reduces precision of the source direction estimate. However, adding blur to the cones does not affect the precision of the direction of the hotspot. In short, by using a fine mesh and adding blur to the cones, the general direction of the source becomes obvious after a just a few reconstructed cones. Then, after a long measurement time, the direction of the source can be estimated with precision but at the expense of a loss in resolution. Thus, if the best angular resolution is required and there are many counts available, thin rings (not blurred by measurement uncertainty) are desired, but for most other situations the standard blurry rings are used.

3.2.5.2 Interaction Separation Distance

Related to the ring width is the selection of a minimum interaction separation distance. This decision comes down to a tradeoff between efficiency and resolution. Since events for which the first two interactions have a small separation distance produce Compton cones with large uncertainties, ignoring these events can drastically improve imaging resolution. However, Figure 3.6 shows the distribution of separation distances between the first two interactions of a 662 keV photon in the 18-detector array described in Section 2.4. Most of the events have a separation of less than 2.5 cm and the distribution peaks at very small separations of less than 0.5 cm. However, there is a set of events where the average interaction separation is about 6 cm. These
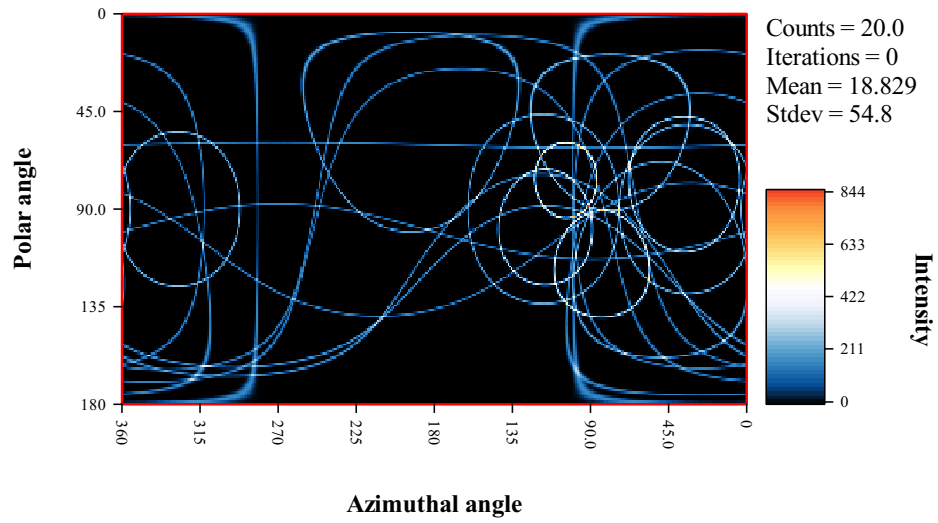
(a)
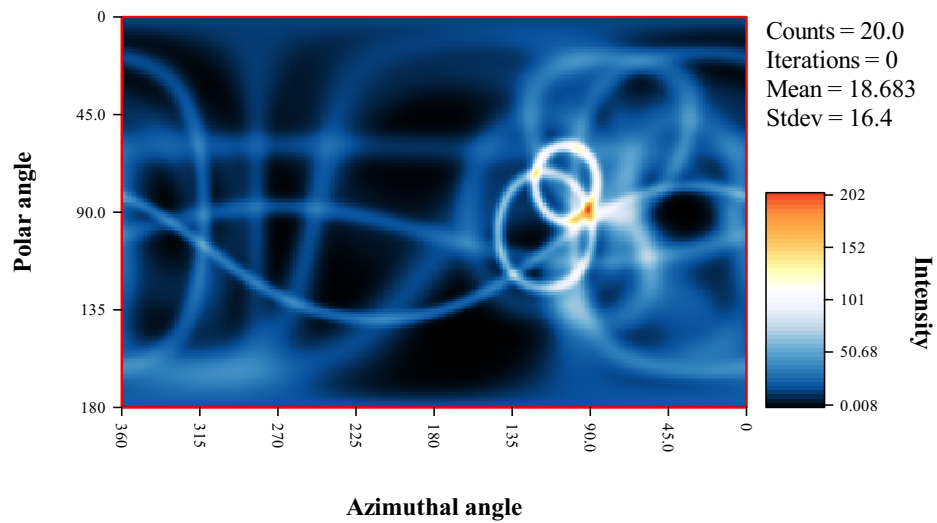
(b)

Figure 3.5: (a) Reconstructed SBP image using 20 thin cones. (b) Reconstructed SBP image using the same 20 cones blurred according the measurement uncertainty.
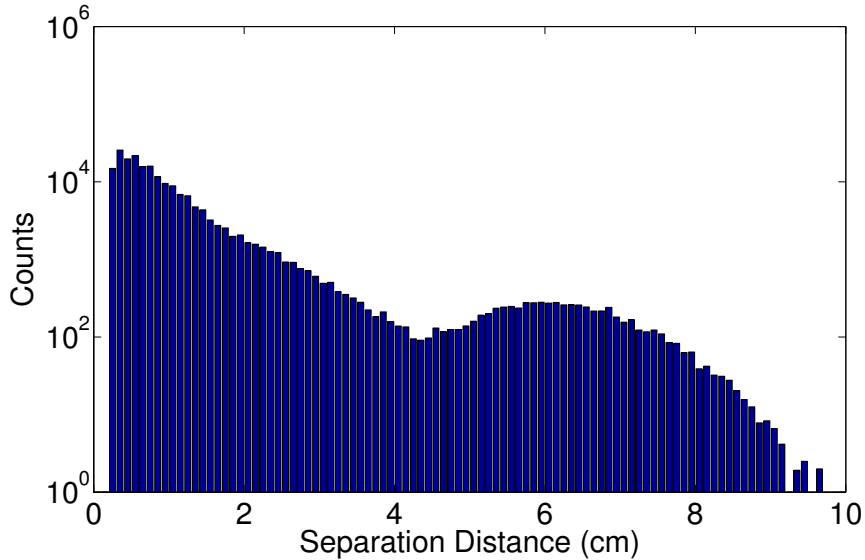
Figure 3.6: The distribution of separation distances between the first two interactions from 662 keV photons in the 18-detector array.

events occur when the incident photon scatters in the front plane of detectors and gets absorbed in the back plane, and we refer to them as "interplane" events. These events are relatively rare happening only about 4% of the time for multiple-pixel events recorded from 662 keV incident photons; however they produce images with much better angular resolution from about 30° FWHM using all events to about 10° FHWM for events that scatter between planes. Since one of the drawbacks of the current array system is the slow readout process, it is advantageous to trigger only on interplane events in high count-rate environments. Using this triggering scheme, the system could operate in very high count rate situations while only recording the best possible events for imaging.

3.2.5.3 Cone Normalization

During the reconstruction process, each event is reconstructed in its own individual image then added to the overall image. Thus, there is a chance to normalize the ring before it is added to the image. To normalize the cone, each element of the

individual cone image is divided by the spatial integral of that image. The advantage of normalizing the cones before adding them to the overall image is that each event contributes the same amount of total intensity to the image. It also means that the events that have large separation distances will have a greater impact on the image since they have more concentrated rings (thinner widths). This emphasis of the large separation events helps to improve the resolution since they should have lower uncertainty than the events with relatively close interactions.

There are, however, a few drawbacks to this normalization technique. For the same reason that the large separation events have greater influence, so do events that have scatter angles near 0° and 180° since they are also concentrated in a smaller area. The difference is that the low and high scatter angle events do not necessarily have smaller uncertainty.

However, if for some reason the reconstruction imaging domain did not encompass the full 4π space around the detector, the cones must not be normalized. The justification is that if only a small part of the ring lies in the imaging domain, that small part of the ring would be normalized and cause the boundary of the imaging space it intersects to be artificially inflated. As seen in Figure 3.7, the corners of the imaging space are affected the most by ring normalization since that is where it is most likely that a small fraction of a ring will intersect the space at a single point and be hugely amplified by the normalization process. At the energies of interest for this work, the difference in image quality is almost negligible between normalized and the non-normalized cones when reconstructing on the full 4π imaging space. Thus, although the standard setting is to normalize the rings, not normalizing them does not cause a significant degradation in image quality.
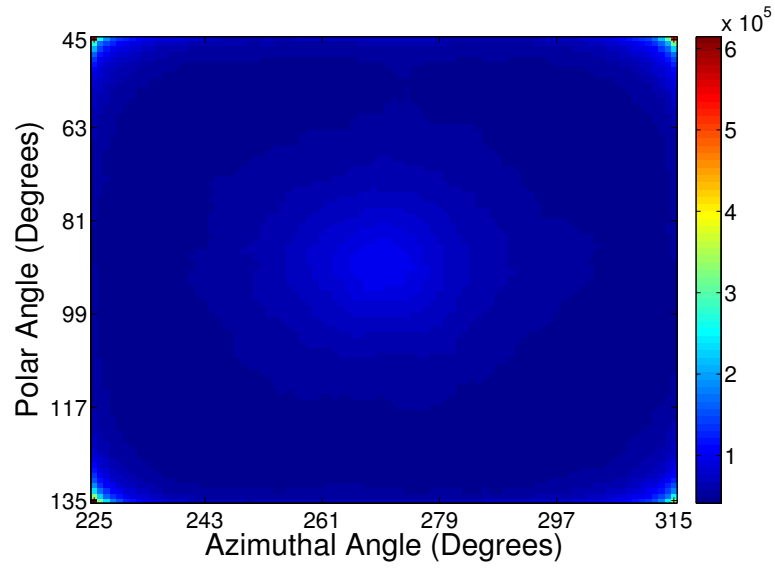
(a)
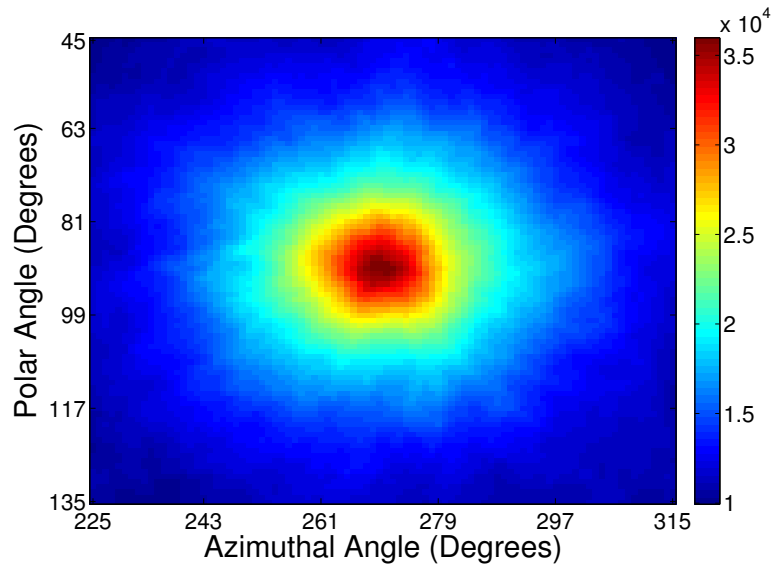
(b)

Figure 3.7: (a) SBP reconstruction using normalized rings on an image space that does not span all 4π. (b) SBP reconstruction using non-normalized rings on an image space that does not span all 4π.

3.2.6 Performance

3.2.6.1 Spatial Resolution

Since this work is focused on reconstructing spatial distributions of point sources in the far field, we define the angular resolution to be the smallest angular distance for which two point sources can be resolved in the image. For SBP, which is a linear algorithm, the imaging resolution can be represented by the full-width at half maximum (FWHM) of the reconstructed spatial distribution. Here, two hotspots in the image that have an angular separation equal to or greater than the FWHM are considered to be fully resolved. Since the imaging space is two dimensional, the FWHM of the azimuthal and polar directions is normally reported.

Figure 3.8 shows the image quality for SBP with the standard settings and options in the UMImaging software (the most important of which is including all interaction-separation distances) for a ^{137}Cs source in the far field in two different positions. These two source positions represent the best and worst directions for imaging performance, respectively, for this system. Notice that the image of the source in the cathode side of the array has a more circular shape with a resolution of about 30° FWHM in the polar and azimuthal directions, and the image of the source in the front side of the array has a more elliptical shape with a resolution of about 30° in the polar direction but almost 55° in the azimuthal direction. This poor resolution in the azimuthal direction is a result of the large uncertainty in the depth of interaction information for multiple-interaction events in the detectors. It is exaggerated for this source position because of the distribution of interaction scatter angles and the influence that the depth uncertainty has as a function of interaction-depth separation. The interaction depth-uncertainty only affects the azimuthal angle of the cone axis as

shown by Xu [5]:

$$\sigma_\theta^2 = \frac{\left[(z_2 - z_1)^2 \cdot p^2/6 + 2[(x_2 - x_1)^2 + (y_2 - y_1)^2](\Delta z)^2\right]}{\left[(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2)\right]^2} \quad (3.6)$$

where σ_θ^2 is the variance of the cone axis in the azimuthal direction, (x_1, y_1, z_1) is the position of the first interaction, (x_2, y_2, z_2) is the position of the second interaction, p is the pixel pitch, and Δz is the interaction-depth uncertainty. Since at 662 keV the probability of forward, side, and backscatter are relatively similar, when the source is on the cathode side, most of the events have large depth separation because most of the events will either backscatter or forward-scatter. Thus, for the average separation distance, the cone-axis uncertainty is more dependent on the pixel pitch instead of the depth uncertainty. However, when the source is on the side of the detector, most of the events have small depth separations because of the large fraction of forward-scatter and backscatter. In this case, now the majority of the cone-axis uncertainty is due to the depth uncertainty. Thus, regardless of the source position, if events with large interaction-depth separation are used for the reconstruction, better resolution is achieved. This result follows naturally from Section 3.2.5.2.

Figure 3.9 shows the SBP reconstruction of the same data used to reconstruct the image in Figure 3.8(a) except that only events that had interactions scatter between planes were reconstructed. Because the cone axis direction has very low uncertainty, the cone widths are small and the resolution of the hotspot is greatly improved from about 30° FWHM to about 10° FWHM. This resolution is the experimental limit to the SBP reconstruction for the current array system. However, if the interaction position uncertainty is improved or specialized detector geometries are implemented, major improvements are still possible.

35658

28533

21409

14284

7160

Intensity

360 315 270 225 180 135 90.0 45.0 0

Azimuthal angle

(a)

0

45.0

90.0

135

180

Polar angle

46232

35875

25517

15159

4801

Intensity

360 315 270 225 180 135 90.0 45.0 0

Azimuthal angle
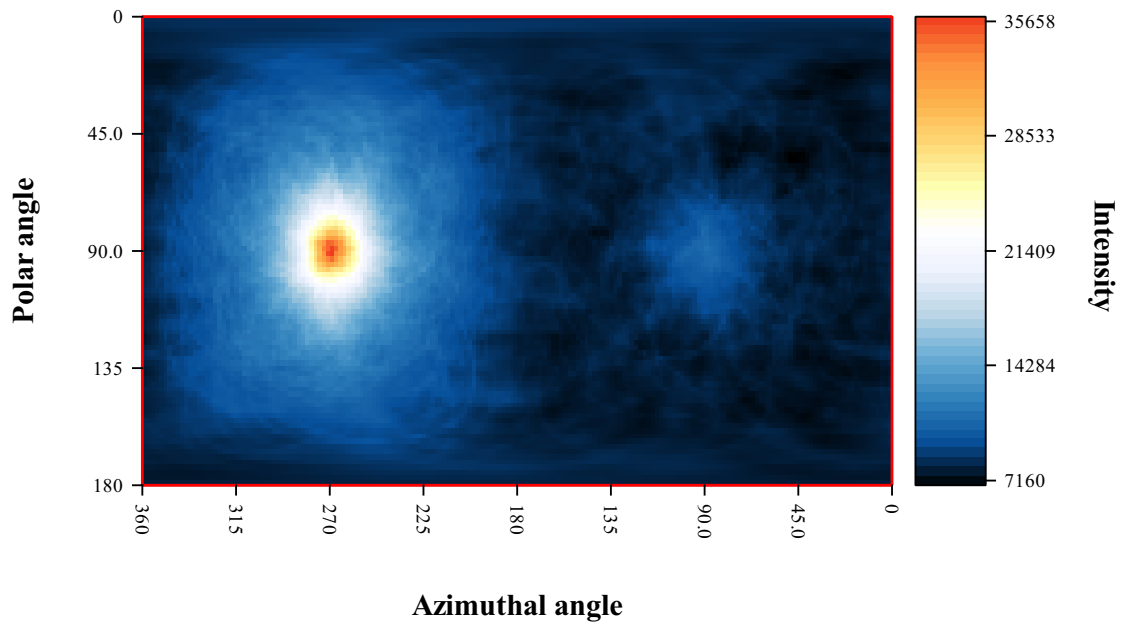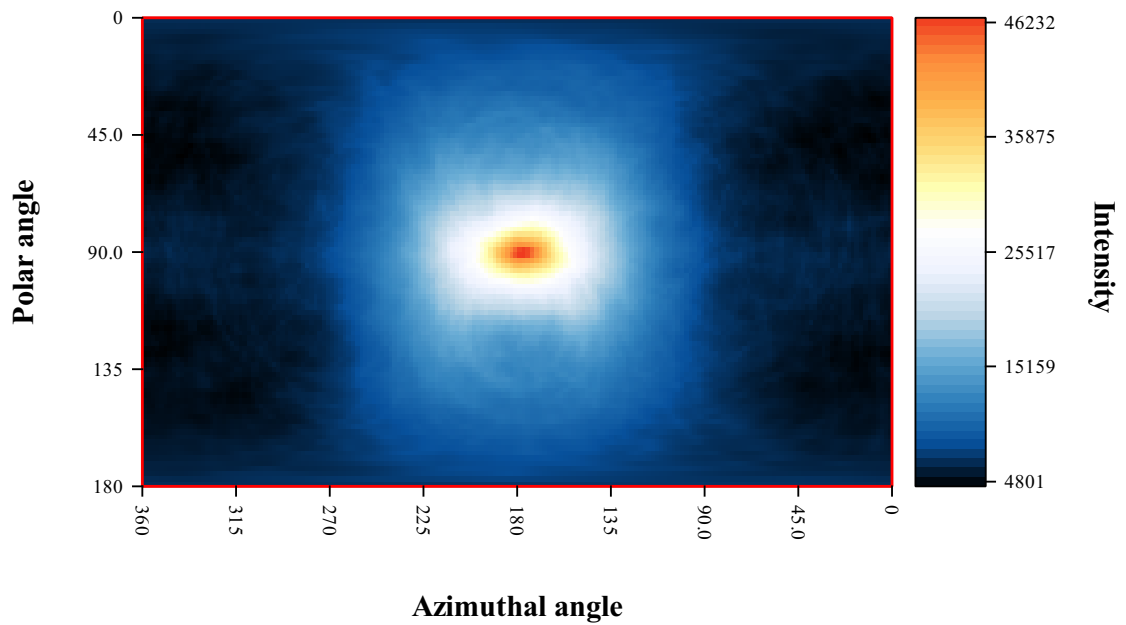
(b)

Figure 3.8: (a) Reconstructed SBP image of a ^{137}Cs source in the cathode direction of the detector array. (b) Reconstructed SBP image of a ^{137}Cs source in the front direction of the detector array.
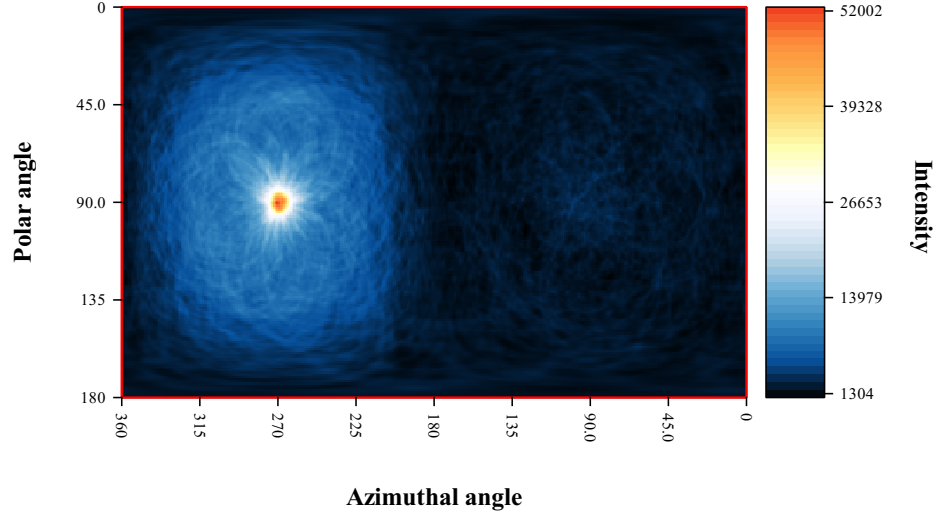
Figure 3.9:
SBP reconstruction of a ^{137}Cs source in cathode direction of the detector array only using events that scattered between the two planes of the detector array.

3.2.6.2 Location Certainty

Even though the SBP reconstruction does not have spectacular spatial resolution, the centroid of the reconstructed hotspot is a very precise measure of the source direction (when using an appropriately fine imaging mesh). After reconstructing about twenty full-energy events, a rough idea of the source direction may be determined, but how many events are needed to know the source direction within a few degrees? To answer this question we look to sample mean and variance.

Consider how an energy spectrum of full-energy deposition events is generated. When an event is recorded, a single count is added to a bin in the distribution. Due to imperfect energy resolution, the location of this bin has some distribution around the true photopeak position. For a sample of n counts, the sample mean is

$$\bar{E} = \frac{1}{n} \sum_{i=1}^{n} E_i. \tag{3.7}$$

The central limit theorem states that $\bar{E}$ converges to the true mean of the photopeak distribution for large sample sets. The standard deviation of the sample mean is equal to

$$\sigma_{\bar{E}} = \frac{\sigma_E}{\sqrt{n}}, \tag{3.8}$$

which implies that the more times you measure some quantity, the less uncertainty you expect in the estimate of its true mean.

One may be tempted to simply apply the same analysis in the imaging domain. However, when a single event is reconstructed using Compton imaging, some intensity is added to many bins in the image space not just one as was the case for the energy spectrum. Thus, the intensities added to the image have some correlation. This correlation is a potential problem since the standard sample statistics assume that the measurements are independent, but the idea that the uncertainty of the estimate decreases with more samples is still valid.

Despite the fact that the assumptions for this analysis are not upheld for Compton image distributions, it is still desirable to test how the source direction uncertainty deviates from that predicted the simple sample statistics. By calculating 100 reconstructions using n simulated ^{137}Cs photopeak events for a variety of values of n, we can directly estimate the value of $\sigma_{\bar{x}}$ where $\bar{x}$ is either the angular distance from the centroid of the hotspot in the reconstructed image or the maximum point in the image to the true source direction.

Figure 3.10 shows $\sigma_{\bar{x}}$ as a function of the number of reconstructed events per trial n for both measures of $\bar{x}$. As expected, $\sigma_{\bar{x}}$ drops as the number of reconstructed evens increases. However, if instead we plot $\sigma_{\bar{x}} \cdot \sqrt{n}$ as a function of n, we can get an estimate of the true standard deviation of the angular distance between the centroid point in the reconstructed image and the true source direction (not the standard

deviation of the sample mean)

$$\sigma_{\bar{x}} \cdot \sqrt{n} = \left(\frac{\sigma_x}{\sqrt{n}}\right) \cdot \sqrt{n} = \sigma_x. \tag{3.9}$$

Figure 3.11 shows the same two distributions as Figure 3.10 but each point is scaled by $\sqrt{n}$. This plot reveals how closely the standard deviation of $\bar{x}$ follows the predicted model since it should be a constant value, σ_x. The distributions for both measures of $\bar{x}$ are in fact flat with mean values of about 50° for the centroid calculation and 20° for the maximum point calculation. The estimate of σ_x using the centroid of the reconstructed hotspot is much higher than the estimate using the maximum point in the image because the centroid is greatly affected by the variance in cone axes and half angles which cause the distribution to be asymmetrical and bias the centroid estimate. The centroid method is also influenced by rings with generated with incorrect sequence reconstruction causing those rings to not pass the source direction at all. Using the maximum point in the image reduces these biases because the symmetry of the hotspot does not affect the location of the maximum point. However, before the hotspot becomes stable at around 100 counts, the maximum point in the image varies greatly and is not a good estimate of the source direction.

Compare either of these estimates of σ_x with the standard deviation of the hotspot in the image (reconstructed with many events) of about 5.7° . This discrepancy between the σ_x estimated with sampling statistics and the standard deviation of the hotspot shape in the image (after many counts) is also a result of the extra variance that the rings add to the reconstruction. However, the takeaway point from this study is that we can estimate the uncertainty of the reconstructed source direction as a function of n. The caveat is that one has to know σ_x beforehand, which is affected by the presence of background, others sources present, and vary as a function of direction. However, for a well characterized system, this estimate of the uncertainty
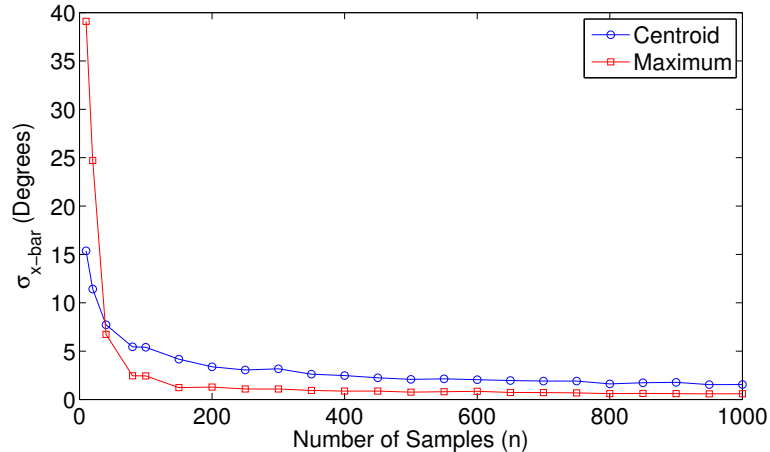
Figure 3.10: Standard deviation of the offset between the true source direction and the centroid or maximum of the SBP image as a function of the number of reconstructed events.

may be useful.

3.3 MLEM Reconstruction

The maximum likelihood expectation maximization (MLEM) algorithm is a widely used and accepted algorithm in the field of gamma-ray imaging [32–34]. In contrast to SBP, which is a linear superposition of the reconstructed cones, MLEM uses the combined information of the entire group of reconstructed events to calculate the most-likely source distribution given the data and a model of the detector system.

3.3.1 Method

In the big picture, there are two parts to the MLEM reconstruction: the likelihood function and the algorithm that solves for the distribution that maximizes that likelihood. The likelihood function contains all of the physics of the problem such as the statistical and probability models. Expectation Maximization (EM) is an iterative algorithm which is used to solve for the source distribution that will maximize
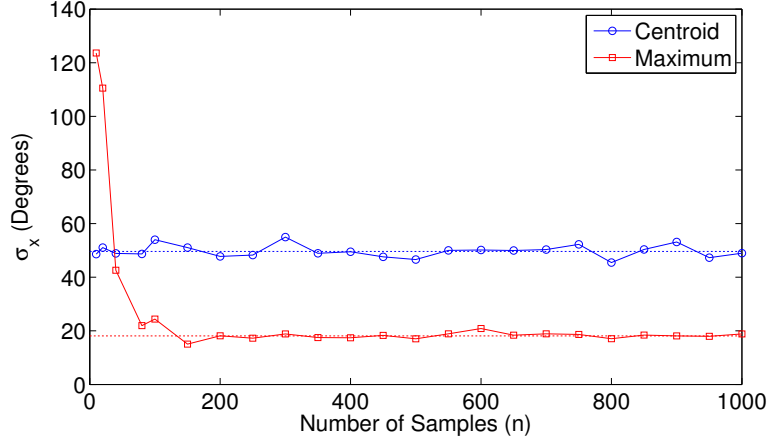
Figure 3.11: Standard deviation of the offset between the true source direction and the centroid or maximum of the SBP image scaled by a factor of $\sqrt{n}$ as a function of the number of reconstructed events (n).

the likelihood function. This iterative algorithm is required to find the most-likely distribution because there is no closed form expression that we can use to solve for it directly as seen in Section 5.3.3.

3.3.1.1 Data Model

Before describing the reconstruction method, we make a brief comparison between binned-mode and list-mode data. Binned-mode data is typical for recording energy spectra or even for imaging techniques like CT or SPECT where the total possible recorded events is a reasonable number with thousands or possibly millions of distinct events that define the set of bins. In this case, it is expected that most bins, which correspond to events with a specific attribute vector $\boldsymbol{v}_m$, will contain many counts. So, for a binned mode reconstruction, it makes sense to store the complete list of possible events and keep a tally of how many of each event type is recorded.

For 3-D position-sensitive CdZnTe detectors used for this work, the number of possible event types is very large. For events that just interact once inside a singe detector, assuming 121 anode pixels, 40 depth bins, and 3000 energy bins, there are $121 \times 40 \times 3000 = 1.452 \times 10^7$ permutations of recorded events. For two interactions

where the pixel numbers are mutually exclusive, the number of possible events is $(121 \times 40 \times 3000) \times (120 \times 40 \times 3000) = 2.09 \times 10^{14}$. One can see how it is impractical to store a complete list of possible events, and even if it were possible almost all of the bins would be empty. Instead, the data is stored as a list of events that were actually detected, which is known as list-mode data.

3.3.1.2 System Model

MLEM is a model-based reconstruction, meaning that is uses the known physical and statistical properties of the system to determine the most likely source distribution. Two key parts of that model are defined here. One component is the sensitivity function

$$s(\vec{r}, E) = \mathsf{P}(D \mid \vec{r}, E) \tag{3.10}$$

which describes the overall probability that a gamma-ray emission from spatial location $\vec{r}$ and energy E is detected (D) by the system regardless of the specific realization of the recorded event.

Another component is the system response function

$$f_m(\vec{r}, E) = \mathsf{p}(\boldsymbol{v}_m \mid \vec{r}, E, D)s(\vec{r}, E) \tag{3.11}$$

which describes the probability of recording $\boldsymbol{v}_m$ for emissions of energy E originating from spatial position $\vec{r}$ for a specific event m given that it is detected. A more detailed description of the first term in (3.11) is described in Section 3.3.2.

3.3.1.3 List-Mode Log Likelihood

Remember that the goal of this method is to reconstruct the gamma-ray source distribution given the set of list-mode data $\boldsymbol{v}$ that was recorded by the detector system. In a general and somewhat abstract form, the likelihood, $\mathcal{L}(\lambda)$, that the

detector system recorded those data during a fixed scan time τ is described by

$$\mathcal{L}(\lambda) = \mathsf{p}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \mid D_1, \ldots, D_n, M = n; \lambda) \mathsf{P}(D_1, \ldots, D_n, M = n; \lambda) \qquad (3.12)$$

where λ is the source intensity distribution in space and energy, D_m denotes that the mth event was detected, M is the random variable associated with the number of recorded counts, and n is the total number of recorded events. Here p is used to describe a probability of a continuous random variable and P is used to describe probabilities of discrete random variables. The first term in (3.12) is the joint probability that each event in $\boldsymbol{v}$ was recorded given that they were all detected and that there were exactly n of them for the true source distribution λ. The second term is then the probability that exactly n events were detected given the true source distribution λ. In short, by finding the λ that maximizes $\mathcal{L}$, one maximizes the likelihood of recording this specific set of recorded events $\boldsymbol{v}$.

Barret and Parra [35] derived a simplified expression for this likelihood function. However, it turns out that it is more convenient to work with the natural logarithm of $\mathcal{L}(\lambda)$ for several reasons, which will become more apparent in Chapter V. The simplified list-mode log-likelihood function is

$$\mathsf{L}(\lambda) \equiv \log[\mathcal{L}(\lambda)] = \sum_{m=1}^{n} \log \left(\iint_{\Omega} f_m(\vec{r}, E) \lambda(\vec{r}, E) \, \mathrm{d}\vec{r} \, \mathrm{d}E \right) - \bar{M}_\lambda \qquad (3.13)$$

where the expected number of recorded counts is

$$\bar{M}_\lambda = \tau \iint_{\Omega} s(\vec{r}, E) \lambda(\vec{r}, E) \, \mathrm{d}\vec{r} \, \mathrm{d}E. \qquad (3.14)$$

3.3.1.4 MLEM Algorithm

The ML solution is defined by a distribution $\hat{\lambda}$ that maximizes (3.13)

$$\hat{\lambda}(\vec{r}, E) = \arg\max_{\lambda} \mathsf{L}(\lambda). \tag{3.15}$$

The initial approach to this problem is to simply take the derivative of the log like-lihood and set it equal to zero. To simplify the problem, we first parameterize the source distribution $\lambda(\vec{r}, E)$

$$\lambda(\vec{r}, E) = \sum_{j=1}^{N} \lambda_j b_j(\vec{r}, E) \tag{3.16}$$

where b_j is the basis function (a spatio-energy voxel covering a small 2-D region of the imaging space over an energy interval), λ_j is the intensity coefficient corresponding to the jth basis function, and N is the total number of voxels. In this parameterization, the basis function is unitless and the coefficients have units of "emissions per unit solid angle per unit energy per unit time."

Substitute (3.16) into (3.13) and rearrange terms

$$\mathsf{L}(\lambda) = \sum_{m=1}^{n} \log \left(\sum_{j=1}^{N} \lambda_j \iint_{\Omega} f_m(\vec{r}, E) b_j(\vec{r}, E) \, \mathrm{d}\vec{r} \mathrm{d}E \right)$$
$$- \tau \sum_{j=1}^{N} \lambda_j \iint_{\Omega} s(\vec{r}, E) b(\vec{r}, E) \, \mathrm{d}\vec{r} \mathrm{d}E. \tag{3.17}$$

Now set the derivative of $\mathsf{L}(\lambda)$ to zero

$$\frac{\partial \mathsf{L}(\lambda)}{\partial \lambda_j} = 0 = \sum_{m=1}^{n} \frac{\iint_{\Omega} f_m(\vec{r}, E) b_j(\vec{r}, E) \, \mathrm{d}\vec{r} \mathrm{d}E}{\sum_{j=1}^{N} \lambda_j \iint_{\Omega} f_m(\vec{r}, E) b_j(\vec{r}, E) \, \mathrm{d}\vec{r} \mathrm{d}E}$$
$$- \tau \iint_{\Omega} s(\vec{r}, E) b_j(\vec{r}, E) \, \mathrm{d}\vec{r} \mathrm{d}E. \tag{3.18}$$

46

Unfortunately, this is a non-linear equation and is difficult to solve directly. Parra et al. [36] derived the EM algorithm for this problem which allows one to iteratively solve for the distribution $\hat{\lambda}$ that maximizes the likelihood function

$$\lambda_j^{i+1} = \frac{\lambda_j^i}{s_j} \sum_{m=1}^{n} \frac{f_{jm}}{\sum_{j'=1}^{N} \lambda_{j'}^i f_{j'm}} \tag{3.19}$$

where

$$f_{jm} = \iint_{\Omega} f_m(\vec{r}, E) b_j(\vec{r}, E) \, d\vec{r} dE \tag{3.20}$$

$$s_j = \tau \iint_{\Omega} s(\vec{r}, E) b_j(\vec{r}, E) \, d\vec{r} dE. \tag{3.21}$$

Xu [5] showed that this algorithm has several favorable properties. First, it is a conservative method which means that, independent of the initial estimate of λ, the total number of counts estimated to be detected is equal to the number of measured photons. Also, the algorithm preserves the positivity of the solution. As long as the initial estimate of λ is positive, all subsequent estimates will be positive unless f_m is zero. Finally, the solution is gauranteed to converge to a global maximum, but the converged solution is not unique unless there are more non-redundant measured events than there are parameters in the estimated source distribution.

3.3.2 System Model

The choice of system model determines the physics that are included in the system response function $f_m(\vec{r}, E)$. An ideal model would account for all of the physical effects that govern not only the gamma-ray interactions but also the generation and transport of electric carriers in the detector material and even the response of the read-out electronics. Ideally, the data input into the MLEM reconstruction would

be the raw signal amplitudes measured by the detector system. However, this model would be incredibly complex and difficult to implement. Instead, a separate algorithm [37, 38] (independent of this work) is used to reconstruct the raw output signals into the gamma-ray interaction data in $\boldsymbol{v}$, which is assumed to be the recorded data here.

Even when only considering the physics of the gamma-ray transport in the system model, there is still a range of complexity that can be chosen from to meet the needs of a specific application (or computational power available). Xu [39] originally derived the system response function for the pixelated 3-D CdZeTe detectors used in this work for two-interaction events only. Later, Wang et al. [40] derived the system response function for events consisting of any number of interactions. For simplicity, two-interaction events are used as examples of the following forms of the system response.

This response can be computationally expensive (as seen below) and is required for each iteration of the algorithm. For binned-mode reconstructions, the entire system matrix could be pre-calculated and stored to drastically improve the reconstruction time. However, that is not possible for many Compton imaging systems, and instead of pre-calculating the system response for all types of events, we are required to calculate it for the initial iteration and then store it in memory for quick access for later iterations if possible. The collection of all of these individual responses is known as the "system matrix," where the responses of each event make up the rows of the matrix. This term is used to describe the collection of system response functions even when it is not physically stored.

3.3.2.1 Calculation Overview

Before introducing the specific implementations of the different forms of the system response function, we describe the form of (3.11). Although at first glance it appears that we calculate the system response by finding the conditional probability

of recording an event multiplied by the sensitivity, in reality we calculate the joint probability that the photon is detected and that it originated from a specific energy and direction. By Bayes' rule

$$\mathsf{p}(\boldsymbol{v}, D \mid \vec{r}, E) = \mathsf{p}(\boldsymbol{v} \mid \vec{r}, E, D)\mathsf{P}(D \mid \vec{r}, E) \tag{3.22}$$

$$\mathsf{p}(\boldsymbol{v} \mid \vec{r}, E, D) = \frac{\mathsf{p}(\boldsymbol{v}, D \mid \vec{r}, E)}{\mathsf{P}(D \mid \vec{r}, E)} = \frac{\mathsf{p}(\boldsymbol{v}, D \mid \vec{r}, E)}{s(\vec{r}, E)}. \tag{3.23}$$

Therefore,

$$f_m(\vec{r}, E) = \frac{\mathsf{p}(\boldsymbol{v}, D \mid \vec{r}, E)}{s(\vec{r}, E)} s(\vec{r}, E), \tag{3.24}$$

and the system response can be calculated without knowing the sensitivity function.

This result is important for two reasons. First, at this point the sensitivity is unknown, and the following expressions for $f_m(\vec{r}, E)$ do not explicitly include the sensitivity. Instead they yield the probability of recording and detecting a given event assuming that it came from a given direction and energy, $\mathsf{p}(\boldsymbol{v}, D \mid \vec{r}, E)$. Also, later we use $f_m(\vec{r}, E)$ to estimate $s(\vec{r}, E)$, and it is important to understand how they are related. For more information, Lingenfelter describes in detail the system response and complications brought about by the conditioning on D [41].

3.3.2.2 Simple

The most basic system response calculation that is required for Compton imaging is the reconstruction of the Compton cones. Typically, the simple response also includes the measurement uncertainty, but it is possible to exclude that effect as well and simply reconstruct cones with a preset cone width. As with SBP, we assume the total energy deposited is the full energy of the incident photon.

This response is described by

$$\left[f_m(\vec{r}, E)\right]_{simple} = \frac{1}{\sigma_\theta} \exp\left[-\frac{(\theta_r - \theta_e)^2}{2\sigma_\theta^2}\right], \tag{3.25}$$

where σ_θ is the uncertainty in the back-projected cone angle described by Xu et al [5] and the other terms are located in Table 3.1. As noted by Wang et al. [40], constant terms have been left out of (3.25)-(3.27) since they would be canceled out in the EM update equation.

The simple response expression describes the agreement between the Compton scatter angle calculated using energies and geometry blurred by a Gaussian distribution with a standard deviation of σ_θ. Thus, it has the same form as the simple-back projection reconstruction. However, a key difference between them is that for the MLEM reconstruction all energetically possible sequences are used in the reconstruction. This strategy is important since it allows the algorithm to (indirectly) determine which is the most-likely sequence in the presence of the other recorded events. Thus, the simple system response is really the sum of (3.25) for both sequences if they are energetically possible.

Figure 3.12 shows the response function for a single two-interaction event using the simple system model. Because both sequences were possible for this event, two rings are present.

3.3.2.3 Attenuation and Scatter Probabilities

The standard model of the system response includes the attenuation and scatter probabilities for each interaction in addition to the probabilities used in the simple model. However, this model still assumes the initial photon deposited its full energy in the detector since the final interaction is modeled exclusively as a photoelectric
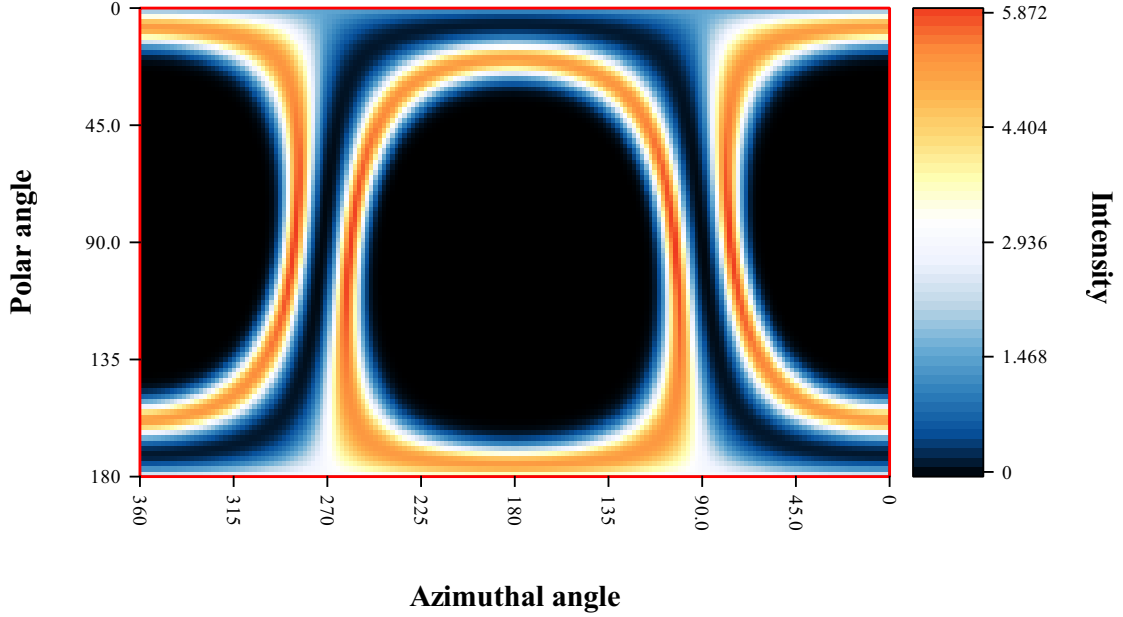
Figure 3.12: Simple system response for a single two-interaction event.

Table 3.1: Symbols used in the system response formulas.

Symbol	Definition
E	Current assumed energy of the incident photon
E_1	Energy deposited in the first interaction
E_2	Energy deposited in the second interaction
$\vec{r}$	Current assumed initial position (direction) of the incident photon
$\vec{r_1}$	Position of the first interaction
$\vec{r_2}$	Position of the second interaction
θ_e	Initial Compton scatter angle calculated using E and E_1
θ_r	Initial Compton scatter angle calculated using $\vec{r}$, $\vec{r_1}$, and $\vec{r_2}$
σ_θ	Uncertainty in the initial Compton scatter angle

absorption. This standard system response takes the following form

$$
\left[f_m(\vec{r}, E)\right]_{standard} = \exp\left[-\int_{\vec{r}}^{\vec{r}_1} \mu_E(\vec{r}') \, \mathrm{d}\vec{r}'\right] \frac{\mathrm{d}\sigma_c(E, E_1, \vec{r}_1)}{\mathrm{d}E'}
$$

$$
\cdot \frac{1}{\sigma_\theta \sin(\theta_e)} \exp\left[-\frac{(\theta_r - \theta_e)^2}{2\sigma_\theta^2}\right] \exp\left[-\int_{\vec{r}_1}^{\vec{r}_2} \mu_{E-E_1}(\vec{r}') \, \mathrm{d}\vec{r}'\right]
$$

$$
\cdot \frac{\mu_p(E - E_1, \vec{r}_2)}{\sqrt{2\pi(\sigma_{E_1}^2 + \sigma_{E_2}^2)}} \exp\left[-\frac{(E - E_1 - E_2)^2}{2(\sigma_{E_1}^2 + \sigma_{E_2}^2)}\right] \qquad (3.26)
$$

where the second term in the first line is the Compton attenuation coefficient [42] and the other variables are again defined in Table 3.1. In physical terms, the first line of (3.26) represents the probability that the incident photon will make it to the first scattering position and undergo a Compton scatter. The second line describes the probability that the scattered photon goes toward the second interaction and travels to the second interaction location. Finally, the last term describes the probability that the scattered photon is absorbed at the assumed energy E within some Gaussian uncertainty.

Figure 3.13 shows the system response of the same event used to demonstrate the simple response shown in Figure 3.12. The addition of the attenuation probabilities is immediately apparent as sections of the reconstructed rings have dramatically less intensity because the incident photon would have had to travel through much more of the detector volume to reach the initial interaction point.

3.3.2.4 Escape Probability

The most complete form of the system model adds the possibility of the photon to scatter and escape the detector without depositing its full energy. This technique
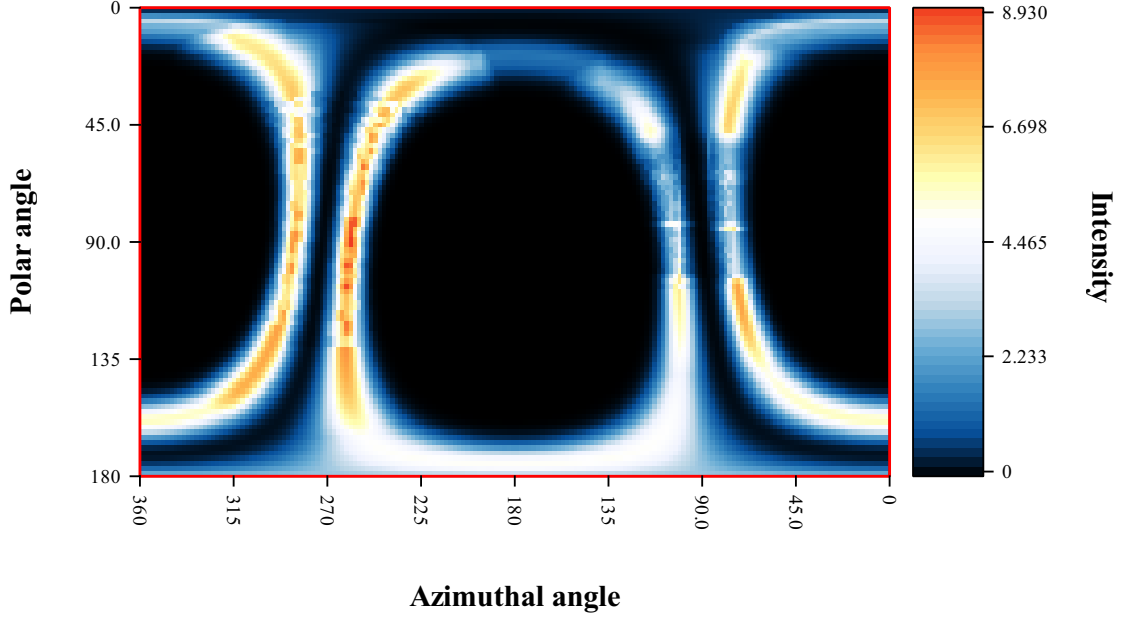
Figure 3.13: Standard system response for a single two-interaction event.

was pioneered by Xu et al. [5], and the response function takes the following form:

$$
\left[f_m(\vec{r}, E)\right]_{complete} = \exp\left[-\int_{\vec{r}}^{\vec{r}_1} \mu_E(\vec{r}\,')\,\mathrm{d}\vec{r}\,'\right] \frac{\mathrm{d}\sigma_c(E, E_1, \vec{r}_1)}{\mathrm{d}E'}
$$

$$
\cdot \frac{1}{\sigma_\theta \sin(\theta_e)} \exp\left[-\frac{(\theta_r - \theta_e)^2}{2\sigma_\theta^2}\right] \exp\left[-\int_{\vec{r}_1}^{\vec{r}_2} \mu_{E-E_1}(\vec{r}\,')\,\mathrm{d}\vec{r}\,'\right]
$$

$$
\cdot \left\{\frac{\mu_p(E - E_1, \vec{r}_2)}{\sqrt{2\pi(\sigma_{E_1}^2 + \sigma_{E_2}^2)}} \exp\left[-\frac{(E - E_1 - E_2)^2}{2(\sigma_{E_1}^2 + \sigma_{E_2}^2)}\right]\right.
$$

$$
\left. + \frac{\mathrm{d}\sigma_c(E - E_1, E_2, \vec{r}_2)}{\mathrm{d}E'} \exp\left[-\int_{\vec{r}_2}^{\infty} \mu_{E-E_1-E_2}(\vec{r}\,')\,\mathrm{d}\vec{r}\,'\right]\right\}. \quad (3.27)
$$

The only difference between (3.26) and (3.27) is the addition of the probability that the photon Compton scatters and is able to escape the detector volume. This small change makes a big difference in the reconstruction. In the previous two system re-

53

sponse models, the energy that a photon deposited in the system was the energy at which the photon was be reconstructed. Now that the model includes the probability of partial energy deposition, a photon that escapes the system can be correctly reconstructed at its true incident energy.

This method has two major implications. One is that more counts can be used for the reconstruction because of the inclusion of partial energy deposition events. The other is that the reconstruction will now estimate the most-likely energy distribution (energy spectrum) for each direction as well as the most-likely spatial distribution for each energy.

Figure 3.14(a) shows the system response in the energy domain of the same event used in the previous two sections for the complete model. Notice that a peak is present at the energy corresponding to the energy that the photon deposited. Above the peak, there is a span of energies from which it is impossible for the photon to have originated because of the scattering kinetics. However, it is possible that the incident photon initially had an energy above that range and scattered out of the system before depositing all of its energy. Figure 3.14(b) shows the spatial response function for the entire energy domain. One can see the two rings from the standard model combined with what is a continuum of rings with smaller scatter angle enclosed in each of the full energy rings. This distribution is expected since the cone axes do not change (these are only dependent on the interaction locations), but the scatter angle gets smaller as the initial energy increases.

3.3.2.5 Sensitivity

The sensitivity of the system needs to be pre-calculated for use in the EM update equation (3.19). In theory, we can calculate the sensitivity function by summing the
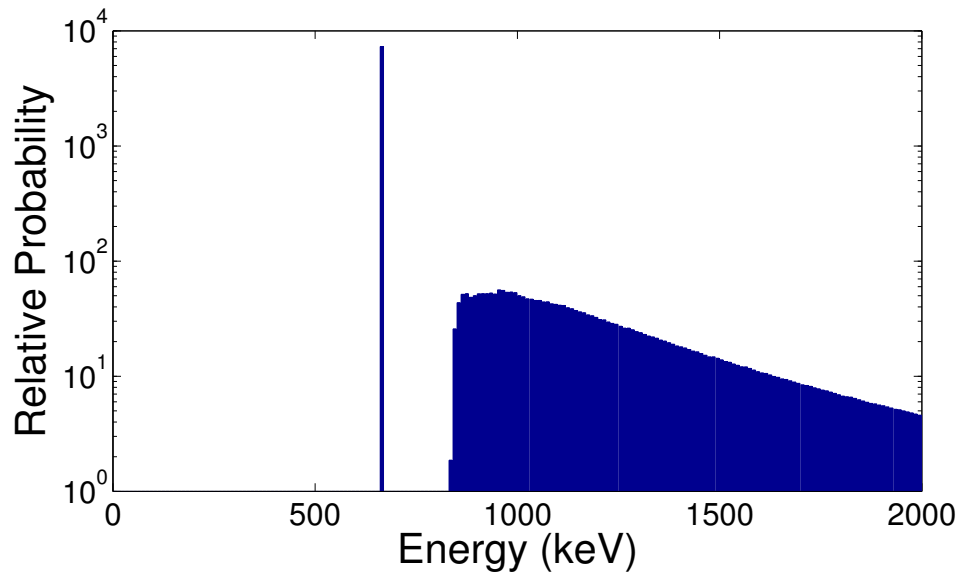
(a)
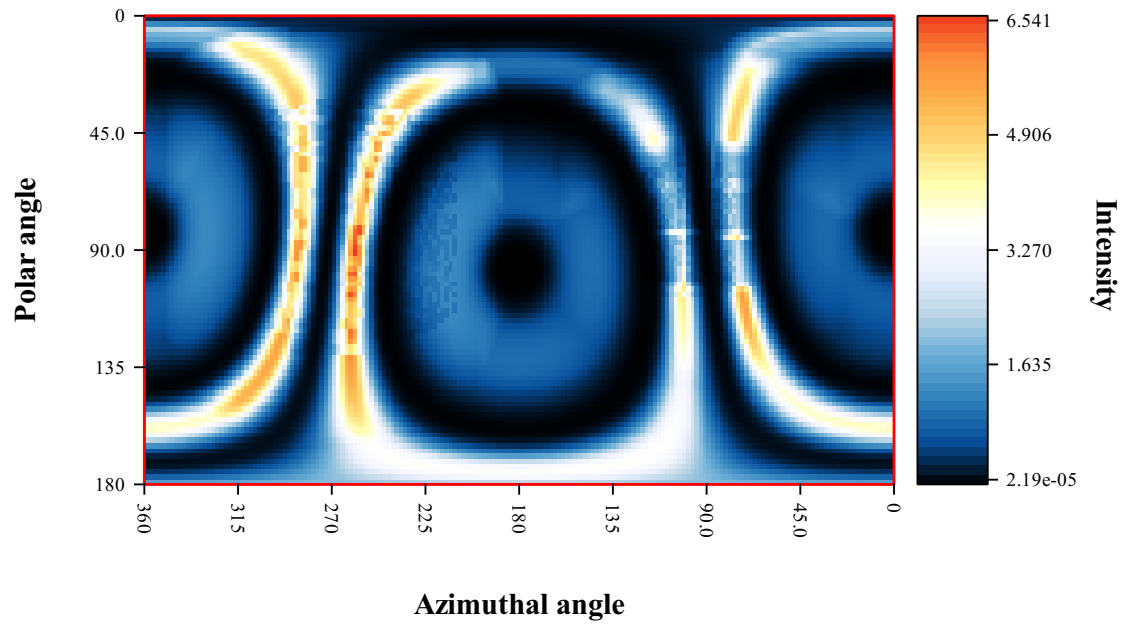
(b)

Figure 3.14: System response for a single two-interaction event using the complete system model in the energy domain (a) and the spatial domain (b).

system response function of every possible recorded event in the system, N,

$$s_j = \sum_{m=1}^{N} f_{jm}.$$ (3.28)

However, there are two problems with this formula. First, we just mentioned in the data model that it would be impossible to even store a list of N events let alone calculate a system response for N events. Also, (3.28) assumes that the the system response is exact, but from the previous section we know we have ignored constant scaling factors and thus (3.28) is not an accurate estimate of the sensitivity.

Xu [5] describes a new way to estimate the sensitivity. In his method the data recorded from a Geant4 simulation of a uniform source in both the spatial and energy domains is reconstructed using the MLEM algorithm with a uniform initial image λ_0. If we perform one iteration of the MLEM algorithm and assume that λ_1 would be identical to λ_0 if the correct sensitivity was used, we can estimate the sensitivity as

$$\lambda_j^1 = \frac{\lambda_j^0}{s_j} \sum_{m=1}^{N_m} \frac{f_{jm}}{\sum_{j'=1}^{N} \lambda_{j'}^0 f_{j'm}}$$ (3.29)

$$s_j = \sum_{m=1}^{N_m} \frac{f_{jm}}{\sum_{j'=1}^{N} f_{j'm}}$$ (3.30)

To simplify the description, Figure 3.15 shows the overall sensitivity of a single detector in the spatial and energy domains (since the sensitivity is more complicated for the 18-detector array). The sensitivity in the energy domain starts out low because of the low probability of Compton scatter in CdZnTe at energies below about 300 keV. The sensitivity then peaks at about 400 keV and drops as the incident photon energy increases. In the imaging domain, the sensitivity is mostly uniform since a single detector is roughly the shape of a cube. However, the largest fluctuations in the distribution are seen at low energies where the surface area dominates the probability of detection. At energies as low as 200 keV, the system is very sensitive to gamma

56

rays incident on the corners of the detector where the most surface area is visible. For moderate energies in the range of 300-1500 keV, surface area and thickness are both important factors and the cathode and anode sides (the poles in Figure 3.15(a)) have the highest sensitivity. At very high energies, the detector sensitivity is very uniform because at high energies surface area and penetration depth become less important and the mass dominates the sensitivity. Thus, for most of the energies of interest, the cathode and anode directions are most sensitive, and that is reflected in Figure 3.15(a).

3.3.3 Performance

3.3.3.1 Resolution

For SBP, the FWHM of the reconstructed spatial distribution of a point source was used to determine how well the algorithm could resolve two point sources. However, since the MLEM reconstruction is not a simple linear operation, the FWHM of a reconstructed point source does not necessarily represent how well the algorithm can separate two nearby point sources. To test this we must actually take a measurement of two point sources separated by a small angle and find the separation at which the sources can no longer be distinguished in the reconstructed image.

For this spatial resolution problem, we use the standard system response model (which assumes full energy deposition) for two reasons. First, if the sources emitted photons with different initial energies, then one would simply have to focus on the individual energy windows that correspond to the energies emitted by each source to separate them in the image domain. Second, the addition of the energy domain in the complete system model adds many more parameters that the reconstruction must estimate. Because we normally have limited data the problem becomes more ill-posed and takes more computing power to solve. This combination of circumstances usually requires us to use a relatively coarse mesh in both the energy and imaging
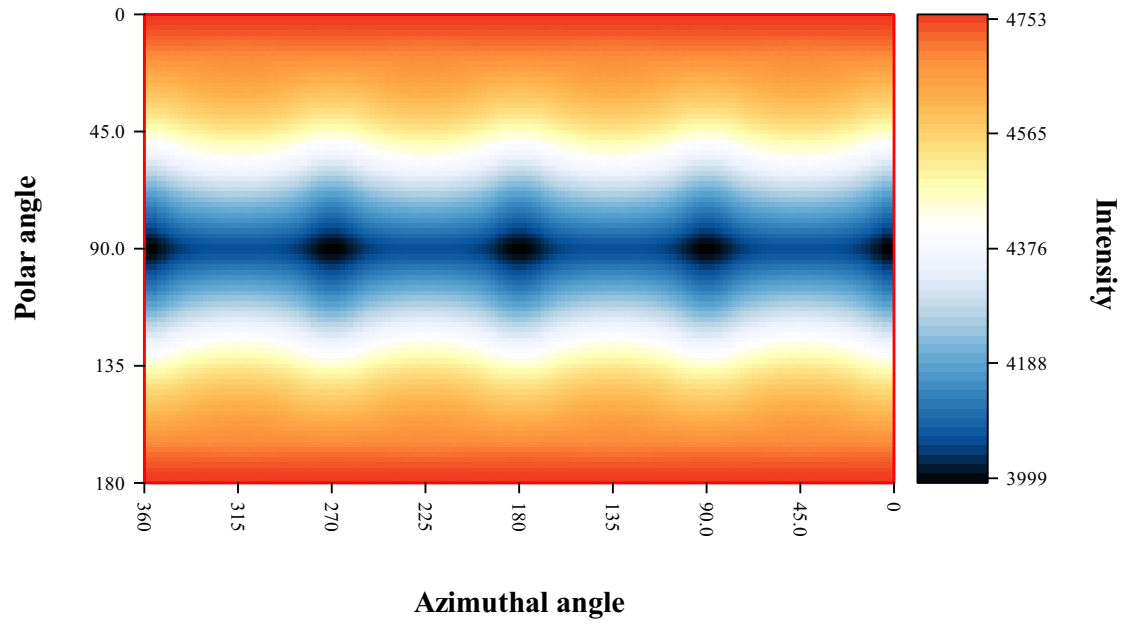
(a)
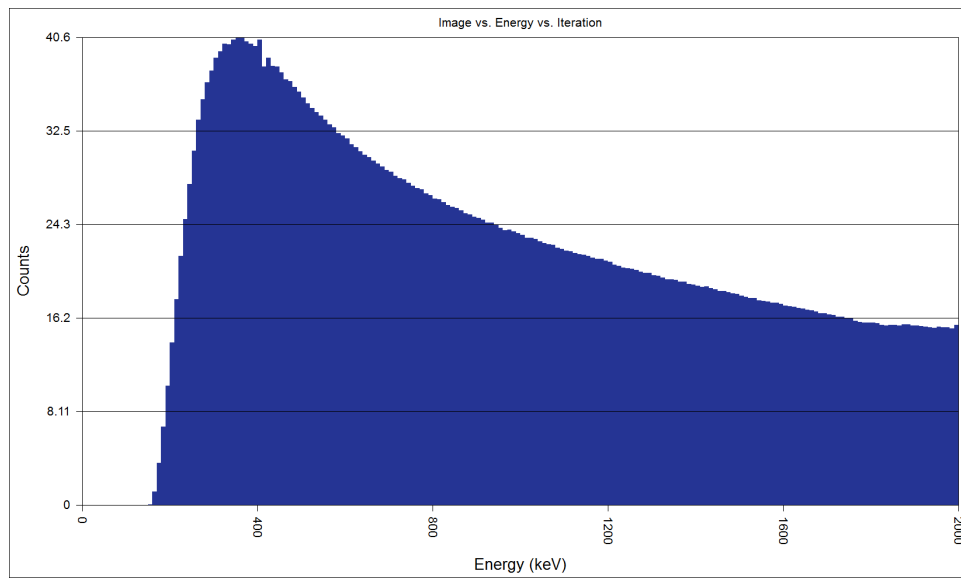
(b)

Figure 3.15: Overall (a) spatial and (b) energy sensitivity of a single detector.

space. Thus, the spatial resolution of the images resulting from the use of the complete system model is typically poor.

To demonstrate the improvement in the spatial resolution between the SBP and MLEM algorithms, we perform an experiment using two ^{137}Cs point sources each with an activity of approximately 30 μCi placed 10° apart in the cathode direction. Figure 3.16 shows the reconstructed image of 10,000 events using SBP and MLEM for a 180 × 360 imaging mesh. The MLEM solution easily separates the two sources, but the SBP solution does not fully resolve the two sources. This result is expected since we showed earlier that the FWHM of the SBP reconstruction is 30° and the sources are closer than that in this experiment. This 10° angular distance is about the limit of the resolution for the MLEM reconstruction using the 18-detector array without requiring long separation distances between interactions (like interpalne events).

For comparison, Figure 3.16(c) shows the reconstructed image using MLEM only using events that have a separation distance of greater than 4 cm. Here the sources are more easily resolved as expected, and when using these select events (which is recorded about once in every 30 ^{137}Cs imageable photopeak events) the best resolution achieved with this system is about 5°.

3.3.3.2 Location Uncertainty

A study similar to that in Section 3.2.6.2 is performed for the MLEM reconstruction. We use the same simulated ^{137}Cs point source data and test procedure for this study that we used test the SBP performance. One notable difference is that the imaging mesh (1° pixel pitch) is more coarse than that used for the SBP reconstruction (0.5° pixel pitch). The purpose of reducing the pixels is to reduce the amount of statistical noise present in the ML solutions. However, Figure 3.18 shows that the estimate σ_x using the centroid of the MLEM image is not constant. Although

59
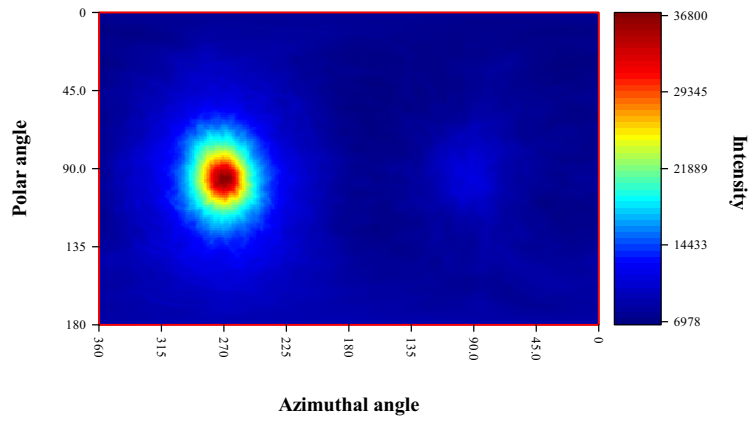
(a)
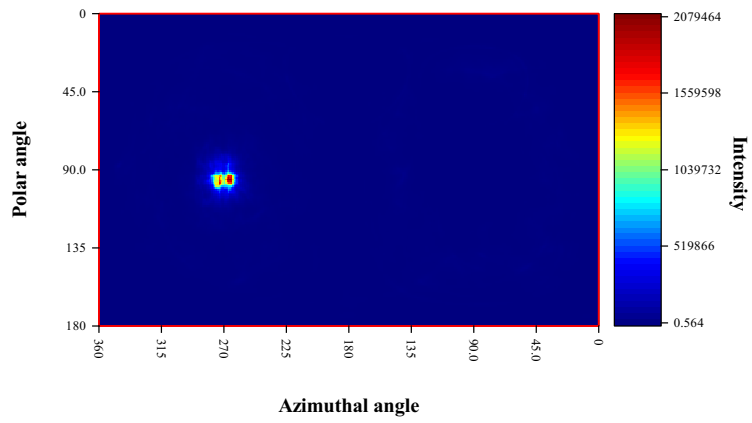
(b)
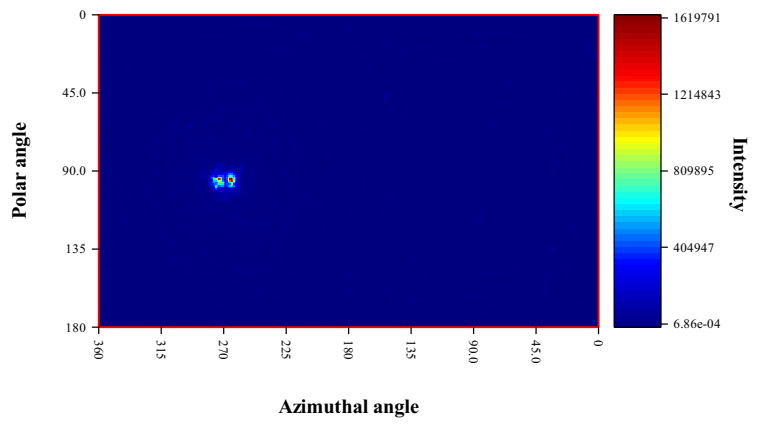
(c)

Figure 3.16: Comparison between the SBP image (a), MLEM image (b), and MLEM interplane image (c) of two ^{137}Cs point sources placed 10° apart in the cathode direction.

60
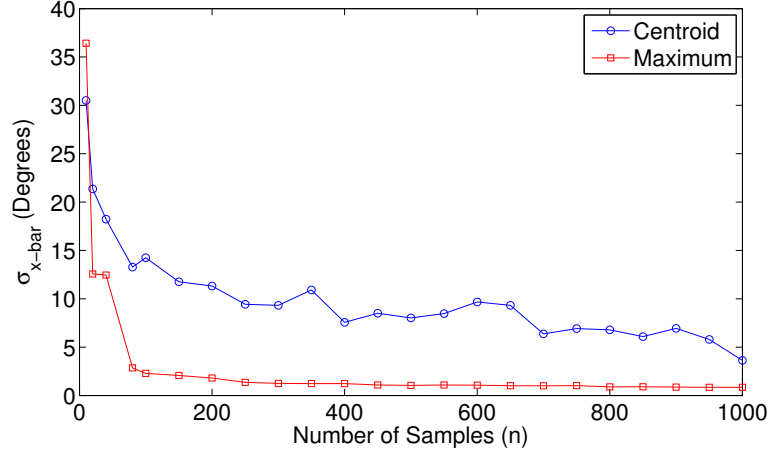
Figure 3.17: Standard deviation of the offset between the true source direction and the centroid or maximum of the MLEM image as a function of the number of reconstructed events, n.

Figure 3.17 shows that $\sigma_{\bar{x}}$ decreases steadily, it does not follow the $1/n$ rate predicted by (3.8). This discrepancy can be attributed to the noise artifacts that are present in the ML solution which causes the centroid to have a larger variance than expected.

On the other hand, the estimate of σ_x using the maximum point in the reconstructed image more closely matches the results using SBP. Figure 3.18 shows that after about a sample size of about 100 counts, the maximum point in the image is consistently in the direction of the source, and the results follow the model in (3.8) yielding an estimate of about $25°$ for σ_x. Here the statistical noise does not effect the maximum point in the image similar to how the asymmetry of the rings does not affect the SBP results that use the maximum point.

These results indicate that although the MLEM reconstruction yields much better angular resolution than SBP, it requires about 1.5 times as many counts to achieve the same uncertainty of SBP when estimating the source direction. This result should be expected because of the tendency the MLEM algorithm has to amplify statistical noise when there are many parameters to estimate (fine imaging mesh) and few reconstructed events.
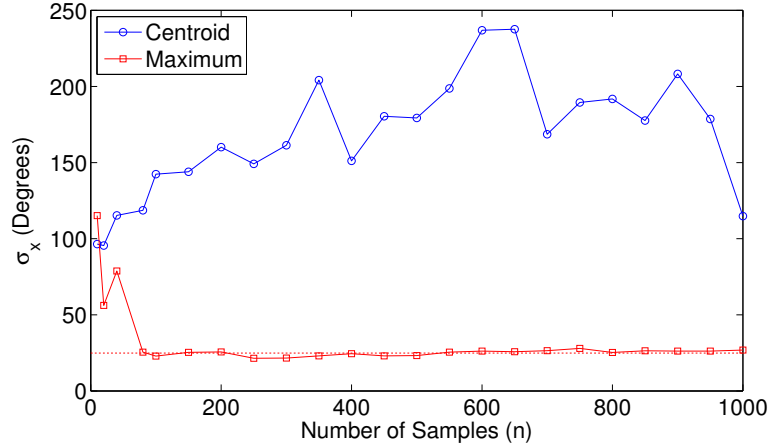
Figure 3.18: Standard deviation of the offset between the true source direction and the centroid or maximum of the MLEM image scaled by a factor of $\sqrt{n}$ as a function of the number of reconstructed events, n.

3.3.3.3   Combined Spectral and Spatial Deconvolution

As mentioned earlier, when using the complete system model, typically the goal is to estimate the incident spectrum from each direction. Even though it is possible to use this full system response to perform high-resolution imaging, computational limitations usually restrict the operator to using a relatively coarse imaging mesh (typically 5° - 10° wide pixels) and energy binning (typically about 4 keV wide bins).

When estimating so many parameters, many reconstructed events are required to reduce the statistical fluctuation in the solution. However, with the large number of voxels in the reconstruction space, the system matrix will be slow to compute and become very large which will quickly fill the computer memory (RAM). These effects tend to compound on each other. As the reconstruction mesh gets finer, more counts are required to reduce statistical fluctuations, and each of those events takes longer to calculate and requires more memory to store. Therefore, the complete model is mostly used to estimate the true incident spectrum deconvolved from the system response and other source directions.

To demonstrate the capability of the MLEM algorithm using the complete system response function, we reconstruct 20,000 events from a measurement of three point sources, ^{137}Cs , ^{22}Na , and ^{60}Co placed in different directions around the array using a reconstruction mesh of 18×36 imaging pixels and 275 energy bins over a range of 300-1400 keV. Figure 3.19 shows the raw mulitple-interaction recorded spectrum and the reconstructed directional spectra. Other than some statistical noise, the directional spectra are reconstructed as the true source emission spectra. Notice that there is no cross contamination in the spectra between from the difference sources (*e.g.,* there is no 511 keV peak in the ^{137}Cs directional spectrum). Also, the ratios of the peaks for the ^{60}Co and the ^{22}Na sources match the true ratios. The ^{60}Co peaks are the same height, and the 511 keV peak in the ^{22}Na spectrum is twice that of the 1274 keV peak. Figure 3.19(b) shows the directional spectra as reconstructed by SBP for comparison.

Figure 3.20 shows the reconstructed spatial distributions for the energy slices associated with the three sources. Some artifacts can be seen in the images which are a result of statistical noise, but the source directions are accurately represented. Because of the coarse mesh, one cannot precisely estimate the source direction, but the reconstructed hotspots do correspond to the true source directions.
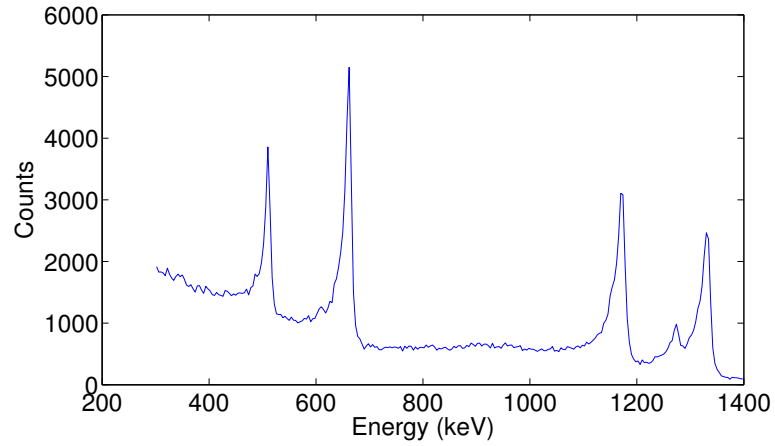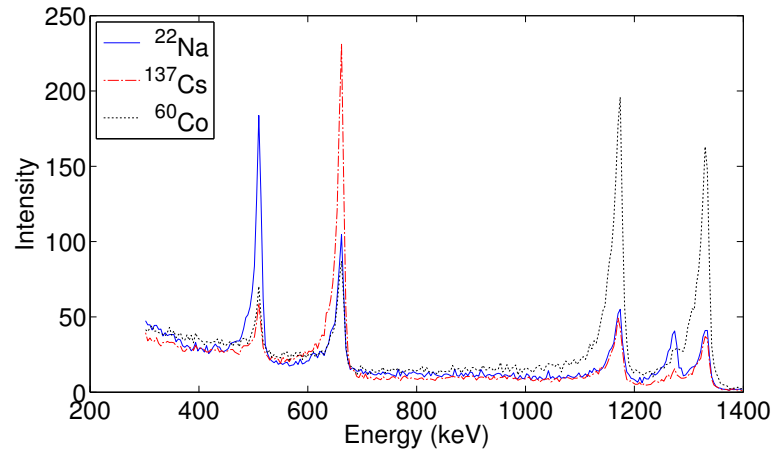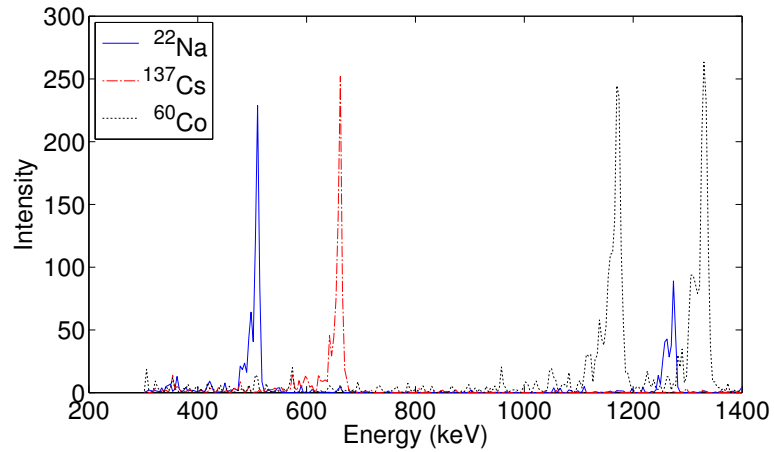
Figure 3.19: (a) Recorded multiple-interaction spectrum for the three-source measurement and the reconstructed spectra as a function of direction for SBP (b) and MLEM using the complete system response (c).

64
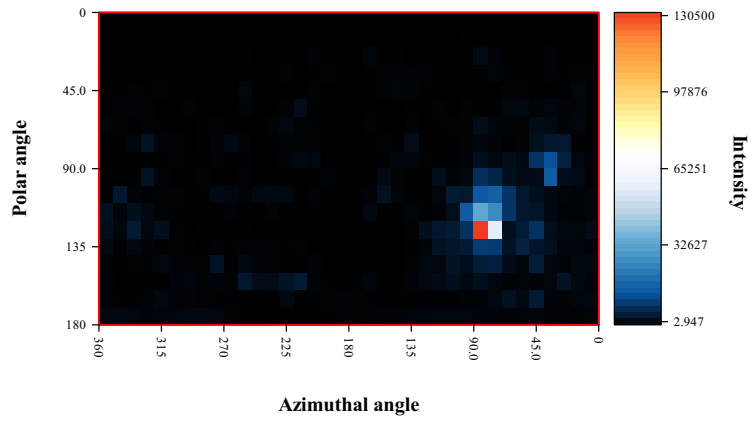
(a)
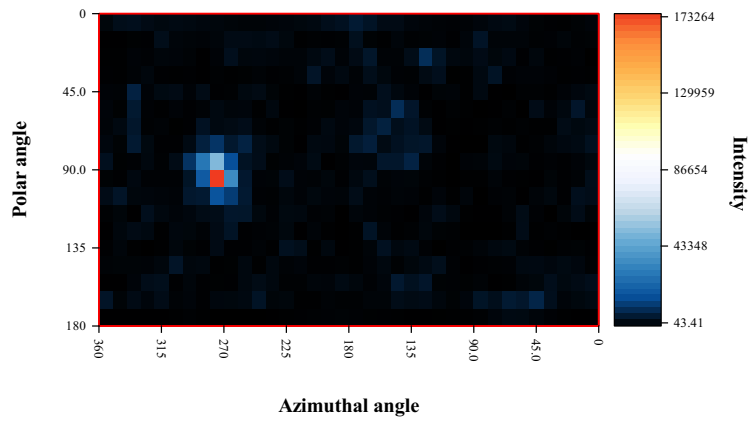
(b)
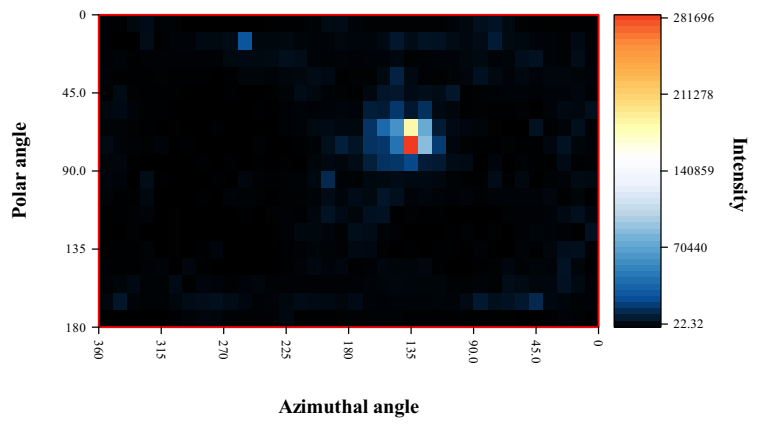
(c)

Figure 3.20: Reconstructed source distributions using the complete MLEM system response for (a) ^{137}Cs , (b) ^{22}Na , and (c) ^{60}Co energy slices.

CHAPTER IV

3-D Image Reconstruction

The reconstruction methods described in the previous chapter focused on far-field 2-D image reconstruction. Although these methods are incredibly useful for finding the direction of a source of gamma-rays, sometimes knowing the actual 3-D location of the source is also important. Especially in complicated environments, the source direction alone may not reveal the object containing the source in question.

There are several related techniques and detector configurations that allow the 3-D localization of sources. This chapter first describes the transition from the 2-D imaging space to the 3-D imaging space. Then it introduces three different types of 3-D imaging which use a single stationary detector array, multiple stationary detector arrays, and a single moving detector array.

For each of the methods introduced, we discuss the changes that must be made to the MLEM system model. However, we use SBP to do much of the analysis because it is convenient and more instructive for describing how the 3-D localization is achieved. SBP is also more convenient to work with since MLEM becomes significantly slower as the size of the mesh is increased for 3-D reconstruction.

Even though 3-D Compton image reconstruction in itself is not a new concept [43, 44], this work will focus on the performance of the 18-detector array for the 3-D reconstructions. Also, the last reconstruction method introduces motion and

time dependence when we demonstrate the ability to estimate the 3-D radiation distribution of a room with a single moving detector array. Another key point is that most of the previous work for 3-D imaging systems is based on traditional Compton camera systems that require interactions to scatter between planes, but in this work the 3-D position sensitivity of the detectors allows a more efficient use of the multiple-interaction events which is especially useful for the moving detector case.

4.1   Comparison to 2D Reconstruction

Before discussing the different forms 3-D image reconstruction, this section highlights a few of the similarities and differences between the 2-D reconstruction and the 3-D reconstruction.

4.1.1   Imaging Domain

Of course, the key difference between 2-D and 3-D reconstructions is the imaging space. For 2-D imaging, the imaging space consists of the far-field imaging sphere represented by spherical coordinates at a large fixed radius. This space essentially consists of a set of possible source directions, whereas the 3-D imaging domain is a volume in the vicinity of the detector array as seen in Figure 4.1. For this work, a standard 3-D Cartesian mesh is used to discretize the volume.

As a result of reconstructing events in a volume, a Compton cone is reconstructed instead of a ring (the intersection of the cone with the 2-D imaging sphere). Figure 4.2 shows a single Compton cone reconstructed in 3D. The cone looks slightly distorted because of the boundaries of the imaging space.

Another effect of the cones being reconstructed into a finite volume, is that normalizing each cone can be problematic for SBP. Just as with a 2-D reconstruction that does not use the entire 4π field of view, if a cone only passes through a small portion of the imaging space, it will have a large relative weight in comparison to cones that
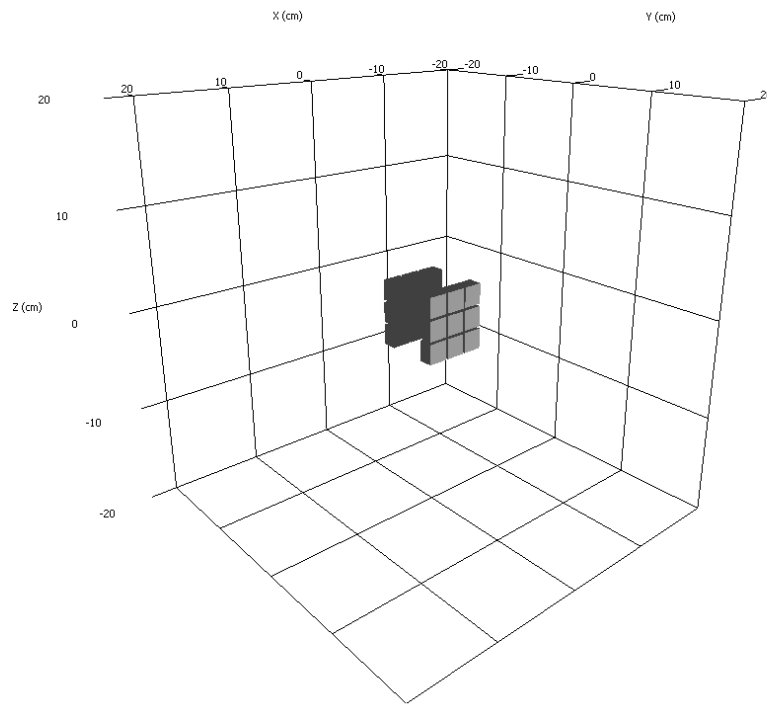
Figure 4.1: 3-D imaging domain. The detector array is located in the center of the image space.
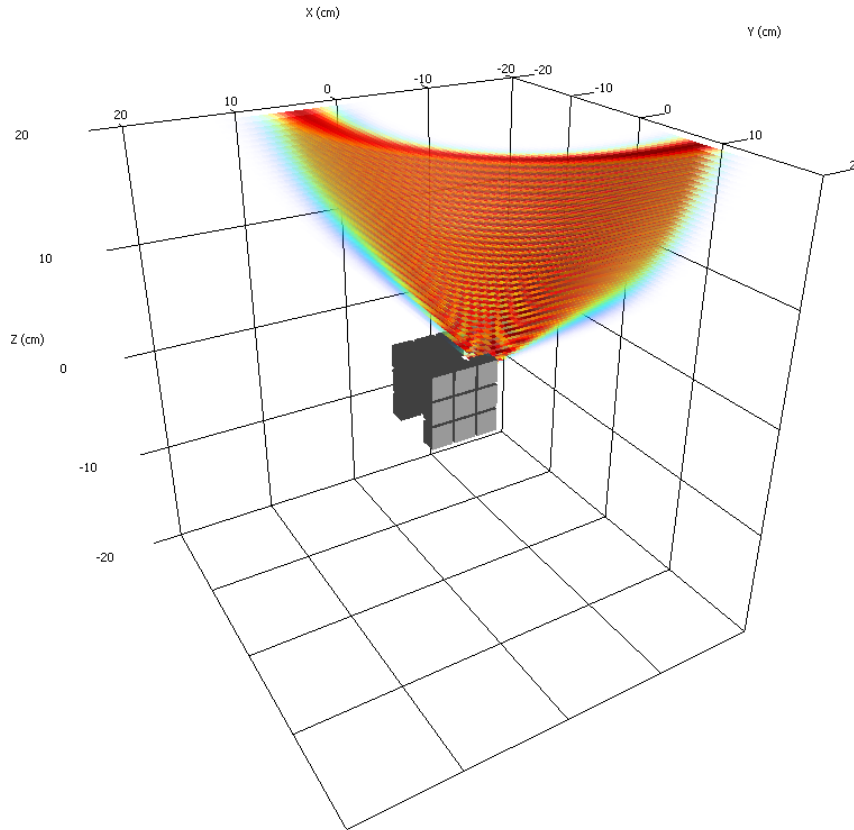
Figure 4.2: A single Compton cone reconstructed in the 3-D imaging domain. Some distortion of the cone can be seen because of the boundaries of the reconstruction space.

pass through a larger fraction of the space. This effect is especially problematic for 3-D imaging because there is always a boundary, unlike 2-D imaging where the entire 4π directional space could be included in the reconstruction. This effect can be mitigated by simply not normalizing the Compton cones or by keeping the detector near the center of the field of view. If the detector is in the center of the reconstruction space, then no matter what direction a cone is reconstructed it will intersect roughly the same number of spatial voxels, assuming the reconstruction space is nearly cubic. However, keeping the detector in the center of the reconstruction space can be wasteful if only a small portion of space on one side of the detector is of interest as a large fraction of computation is used on portions of the space that are unimportant.

4.1.2 Localization

In the 2-D reconstruction space, the direction of a source of incident photons is determined by the intersection of the Compton rings on the imaging sphere. For the 3-D reconstructions discussed below, the ability to localize the position of the source is determined by the amount of parallax present in the measurement.

Parallax is the perceived change in the position of an object due to a change in the viewing position. For example, this effectis apparent when driving in a car. As the car passes nearby trees, the direction of the trees changes dramatically from one position of the car to the next; whereas the mountain in the distance or the sun appear to be stationary no matter how far the car travels. In this case, the position of a tree can be estimated if the position of the car and the relative direction of the tree is known at two (or more) different times. However, since the change in relative direction between the car and the sun is negligible, the 3-D position of the sun cannot be estimated in this way.

Similarly if the vertices of the Compton cones have sufficient parallax in relation to the source, we can estimate the 3-D position of radiation sources. The following sections describe three different ways to achieve this parallax using: a single stationary detector array, multiple stationary arrays, and a single moving detector array.

4.2 Single Array

When only using a single stationary-detector array, the parallax required to localize a source in 3D is achieved by the separation distance of the individual Compton cone vertices inside the array. Figure 4.3(b) demonstrates this concept by showing three Compton cones reconstructed from the 18-detector array. Normally, for what we consider far-field imaging, the actual vertex of the Compton cone is not used in the Compton cone reconstruction. For the 2-D methods described previously, the verte-

ces all of the cones were effectively placed at the origin of the imaging sphere. This simplification made sense for the 2-D case where the imaging sphere has an infinite radius (far field), and the small changes in the vertices of the cones were negligible. Now that the imaging space of interest is much closer to the detector (near field), the differences in the positions of the cone vertices are exactly what allow sources to be located in 3D.

Sullivan et al. [45] described a different way to estimate the source-to-detector distance of point sources using a standard far-field 2-D reconstruction. The idea is that since the source is relatively close to the detector, a blur is created in the reconstructed hotspot because the cones are not projected onto the correct focal plane (similarly to how an optical camera would be out of focus with the wrong focal length). However, this type of analysis only works for single point sources, and the field of view for that system is limited since it uses a traditional Compton camera which requires photons to scatter between two 2-D-position-sensitive detector planes. The same type of algorithm could be implemented for the detectors used in this work, but the 3-D mesh approach allows for the reconstruction of more diverse source distributions (at the expense of more computation).

4.2.1 MLEM Model Changes

Fortunately, the system model was developed in a general way; so we only need to make a couple changes to the 2-D reconstruction to perform the 3-D reconstruction. One new term that needs to be added to the system response functions is the geometrical attenuation from the point of emission to the first interaction location inside the detector volume $\left(4\pi\|\vec{r}_1 - \vec{r}\|^2\right)^{-1}$. The modified simple system response is

$$\left[f_m(\vec{r}, E)\right]_{simple} = \frac{1}{\sigma_\theta\|\vec{r}_1 - \vec{r}\|^2}\exp\left[-\frac{(\theta_r - \theta_e)^2}{2\sigma_\theta^2}\right]. \tag{4.1}$$
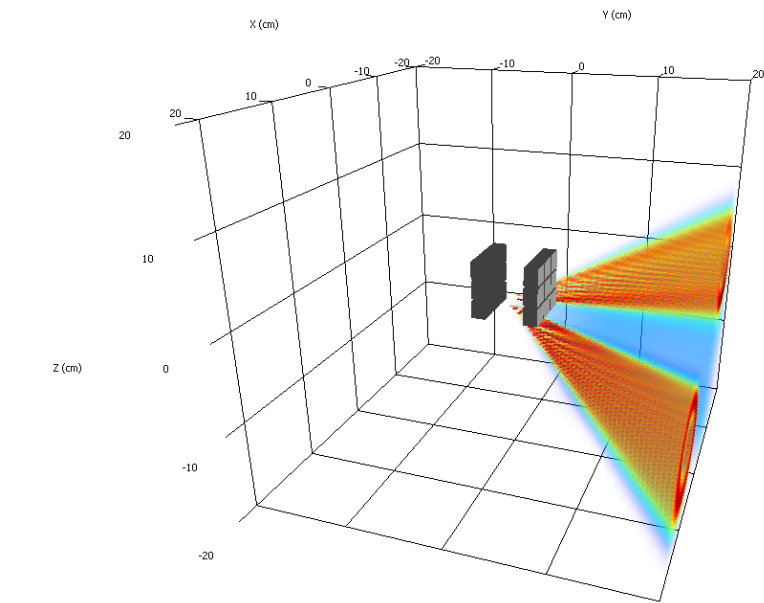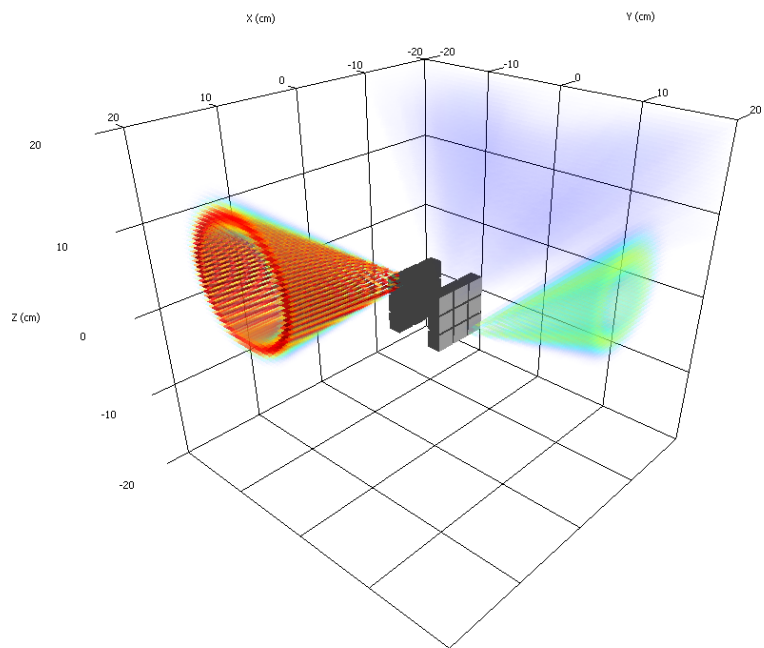
Figure 4.3: (a) Compton cones reconstructed using the far-field assumption where the vertices are all placed at the origin of the detector array. (b) Compton cones reconstructed using the near-field assumption where the vertices are all placed at the location of the first interaction.

The $1/4\pi$ term is excluded because, as a constant, it will cancel out in the EM update equation. The geometrical attenuation was previously ignored because, for the far-field assumption, the distance is constant as well.

Another term that must be altered to correctly model the 3-D case is the sensitivity of the system. Calculating the true sensitivity is still impractical since the number of possible interactions has not been reduced. For the 2-D case, the sensitivity was estimated by performing the first iteration of MLEM (starting with a uniform image) for a large number of emitted photons uniformly distributed in incident direction and energy. A similar process must be followed for the new 3-D space surrounding the detector array. However, now the source in the simulation must be uniformly distributed in the space around the detector with isotropic emissions. Depending on the size of the space and resolution desired, this simulation could take much longer to converge in the spatial domain even though there are still spatial symmetries that can be used to reduce the calculation time up to a factor of eight for a cubic imaging space centered on the detector.

In practice, however, these modifications are not included in the model. For the standard system model (which assumes full energy deposition), the sensitivity is not used for the same reason it is not used in the 2-D case. First, a full energy sensitivity would have to be created for every energy of interest, and second the results without the sensitivity still perform well. Also, the spatial sensitivity does not change rapidly enough to cause problems reconstructing individual point sources, but when using the complete system model including the energy domain, the reconstructed energy spectrum would suffer due to the varying efficiency as a function of energy. To account for the energy dependence, we could use a sensitivity factor that is uniform in space but that has the correct energy dependence. If more complicated source geometries are expected, a more accurate sensitivity estimate may become more important.

4.2.2 Available FOV

The field of view (FOV) of a single array for near-field 3-D image reconstruction is dominated by two factors: the largest separation distance between cone vertices and the measurement uncertainty of the Compton cones. The first factor is determined by the physical extent of the detector perpendicular to the direction of the source, and the second factor can be represented by the angular uncertainty of the array when reconstructing images in 2-D. For this study we will focus on the space directly in front of the cathode surface of a plane of the 18-detector array which has the best angular resolution with the most parallax.

When the source is on the outer edge of the FOV, the most useful Compton cones are those for which the first interaction occurs near the corners of the detector plane because they have the most parallax. Therefore, if we consider the just the events that have first interactions at corners of the plane, we can estimate the limit of the ability to resolve sources as a function of the perpendicular distance to the detector plane surface and the angular resolution of the detector array. This limiting distance then defines an estimate of the region of space surrounding the detector that can be resolved.

Figure 4.4 shows a diagram of what the simple back projection would look like after many reconstructed counts if only events where the first interaction occurred in a corner of the detector array are considered. Although it appears that the cones depicted in the figure are Compton cones, they in fact represent the surfaces of solid cones formed by the summation of many Compton cones emitted from the two corner voxels. The intersection of one of these cones with the 2-D imaging sphere would be analogous to the hotspot in the source direction of a 2-D reconstruction.

In the diagram, d is the distance from the surface of the detector array to the source, a is the distance from the center of the array to the corner pixel of the array, α is the angle between the detector surface and the source position measured at the

corner pixels, and σ_α is the uncertainty in α. The uncertainty σ_α (twice the half-angle of the cone) can be represented by the FWHM of the hotspot from 2-D reconstruction. σ_d is then the FWHM of the estimated distance from the source to the surface of the detector array which can be calculated using the following formula

$$\sigma_d = \left(d + \frac{\sigma_d}{2} \right) - \left(d - \frac{\sigma_d}{2} \right)$$
$$= a \cdot \tan \left(\alpha + \frac{\sigma_\alpha}{2} \right) - a \cdot \tan \left(\alpha - \frac{\sigma_\alpha}{2} \right)$$
$$\frac{\sigma_d}{a} = \tan \left(\alpha + \frac{\sigma_\alpha}{2} \right) - \tan \left(\alpha - \frac{\sigma_\alpha}{2} \right) \tag{4.2}$$

where

$$\alpha = \tan^{-1} \left(\frac{d}{a} \right). \tag{4.3}$$

Note that (4.2) is only valid when estimating the FWHM for large distances d. When the source is very close to the detector (on the order of a detector-pixel pitch), the edge pixels will play a less important role than the pixels near the center of the detector.

Figure 4.5(a) shows the trends predicted by (4.2) and the measured trends from a reconstruction using simulated data from the standard detector array and an angular resolution of 5°. Notice that the simulated results do not agree with the calculated results, but they are scaled by a factor of two. We expect that the result should not follow (4.2) exactly because only a fraction of the reconstructed cones have vertices near the edges of the detector. Many of the cones are backprojected from the center of the array which would substantially degrade the resolution.

Unfortunately, as the angular resolution of the detecor degrades, the FWHM of the distance estimate deviates greatly from that predicted by (4.2). This result indicates that for detectors with good angular resolution, the cones projected from the edges of the array dominate the ability to resolve the source distance. As the resolution
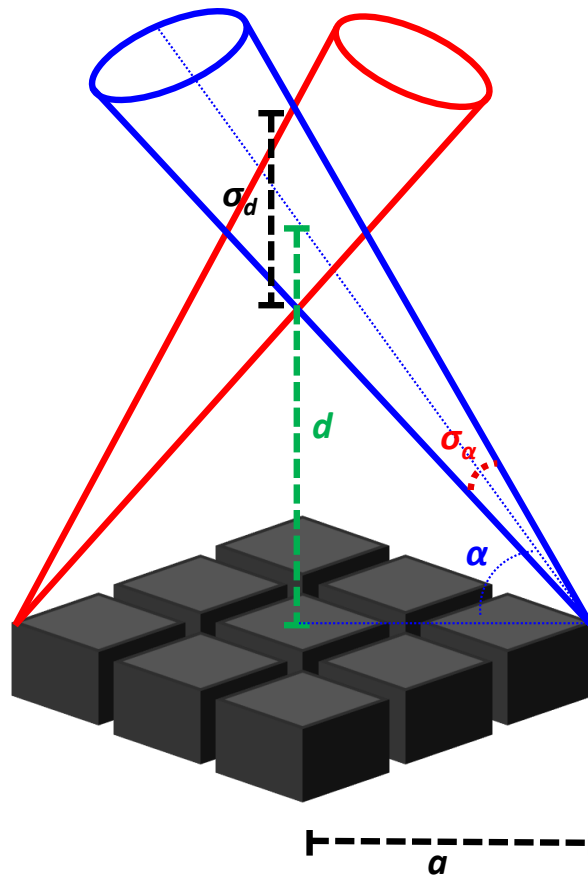
Figure 4.4: Diagram of the extreme case where only events where the first interaction occurred in the corner detector-pixels are used for near-field 3-D imaging. The cones depicted here represent the surfaces of solid cones formed by the summation of many Compton cones originating form the two corner pixels.

degrades, the cones projected from the center area of the array add an increasing amount of uncertainty in the distance estimate. Thus, the exclusive use of cones originating from the corner pixels becomes an increasingly poor assumption. Also, as the detector size increases, the fraction of cones projected from the edge pixels decreases, making (4.2) an even poorer estimation of the distance resolution.

Therefore, the angular resolution of the detector is the most important factor when resolving the source distance. For detectors with excellent angular resolution, (4.2) predicts a source-distance resolution that is proportional to the true resolution for a given detector size. This constant of proportionallity can be easily estimated with a single measurement or simulation. Alternatively, we could use (4.2) to estimate the required detector size to achieve a given source-distance resolution.

4.2.3 Localization Uncertainty

Figure 4.6 shows the reconstructed image of a simulated ^{137}Cs point source placed directly above one plane of the standard detector array using the SBP algorithm for events that scatter between planes. In this simple example, the point source is reconstructed almost exactly at the true source position 10 cm from the cathode surface.

These are atypical results for the 3-D SBP reconstruction. For data with more uncertainty (*e.g.,* measured data) the centroid of the reconstructed hotspot is biased toward distances farther from the detector than the true source location. This bias can be seen in Figure 4.4 as the diamond shape region where the two cones overlaps has a bias away from the detector plane. When all imageable events from the simulation used to generate the results seen in Figure 4.6 are reconstructed (not just interplane events), the center of the hotspot is nearly 11 cm from the center of the array instead of 10 cm where the source was in the simulation. Experimental data shows the same behavior except further exaggerated by relatively poor angular resolution. Thus, the
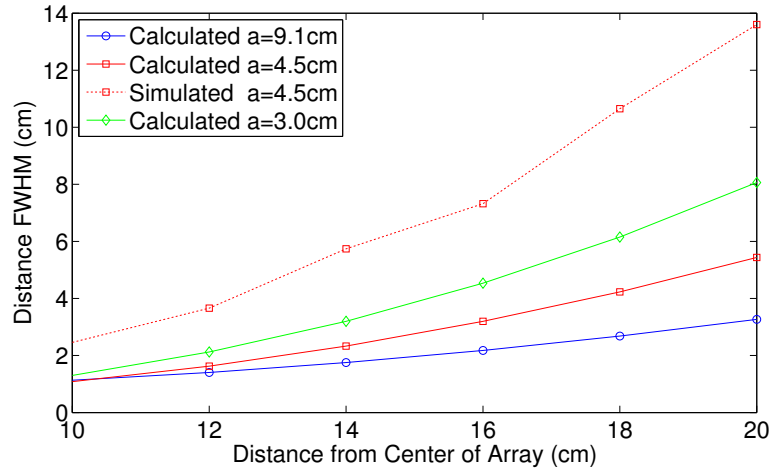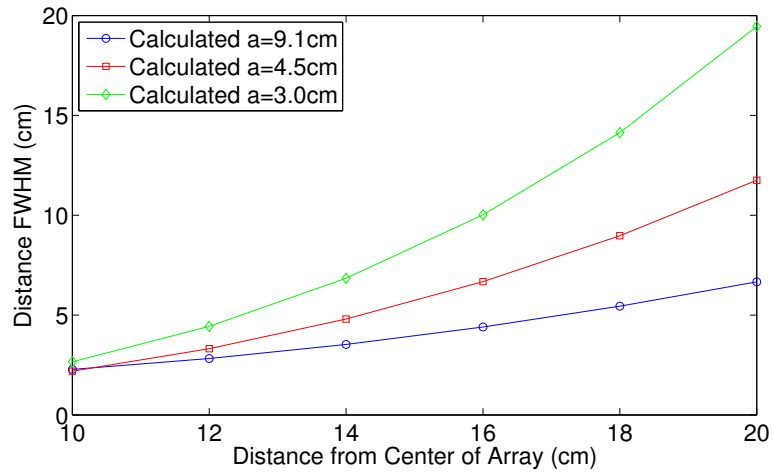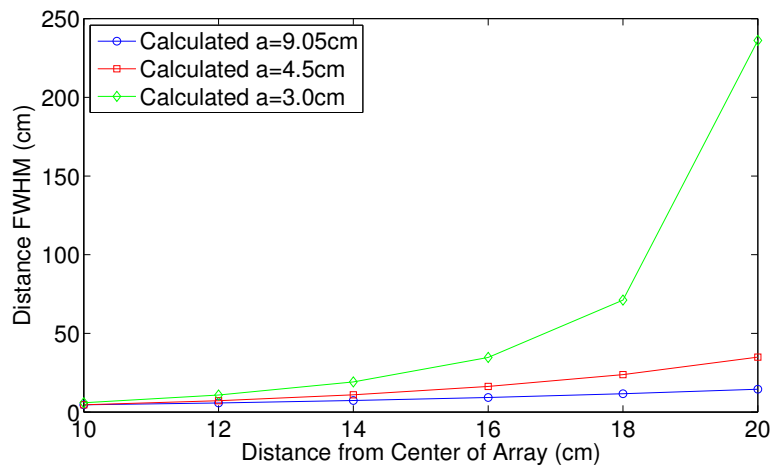
Figure 4.5: The FWHM of the source to detector distance estimate as a function of the true source to detector distance and the detector size for an angular uncertainty of: (a) 5° , (b) 10° , and (c) 20° .
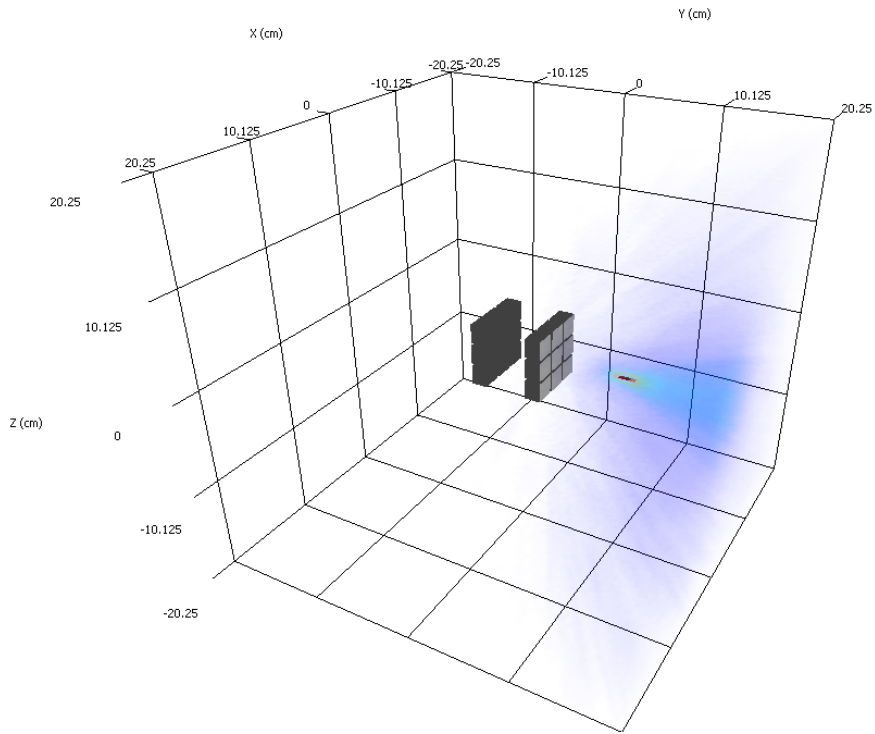
(a)
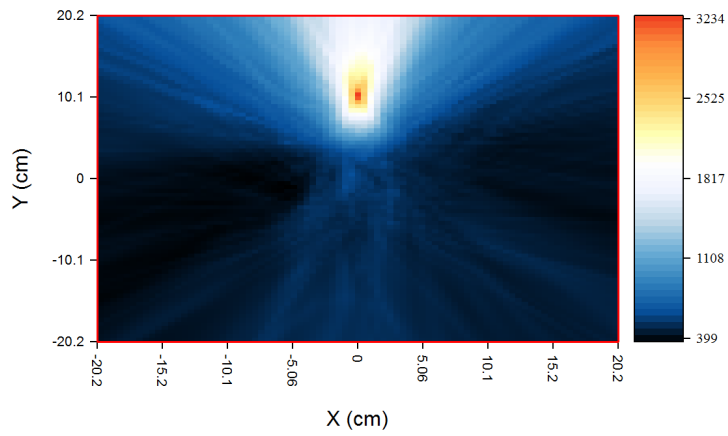
(b)

Figure 4.6: (a) 3-D rendering of the SBP image reconstruction of a ^{137}Cs point source placed 10 cm from the center of the detector array using only interplane events. (b) X-Y slice of the reconstructed image shown in (a).

angular resolution is of utmost importance for estimating the source position (which is related to the ability to resolve sources at different distances).

If we account for the offset in the estimate of the centroid of the source position, we can perform a similar analysis as in Section 3.2.6.2 to determine if the 3-D reconstruction follows a similar reduction in location uncertainty as a function of reconstructed events. The simulated ^{137}Cs photopeak data at a distance of 10 cm is used for this study (without a requirement to scatter between planes). Figure 4.7(b) shows the standard deviation of the offset between the true source position (shifted by the systematic offset) and the centroid or maximum of the hotspot in the image scaled by $\sqrt{n}$ as a function of the number of reconstructed events (n). Again, the important result is that σ_x is constant as a function of the number of samples, which means that it follows our expectations and the uncertainty in the estimate of the source location can be approximated by (3.8). Of course, for that formula to be useful, σ_x must be known for the specific scenario, and for the 3-D case it is likely to be a strong function of the source distance and the angular resolution of the detector. Also, notice that in this case, there is little difference between the estimate of σ_x using the centroid of the image and the maximum in the image. This is a result of the added dimension which makes the maximum point vary more and the fact that the imaging space is centered on the source location and does not extend far past the source location. Thus, the centroid is affected less by the extra variance of the cones.

4.3   Multiple Arrays

Another form of 3-D imaging can be achieved by using multiple detector array systems where the two arrays are strategically placed near the 3-D region of interest. In this case, the parallax is attained by the offset in the position of the arrays, and thus, the scale of the reconstruction space can potentially be much larger than the dimensions of any individual array.
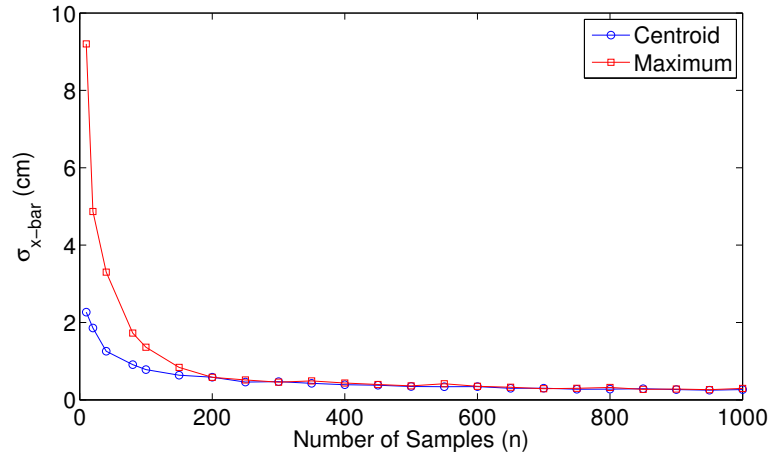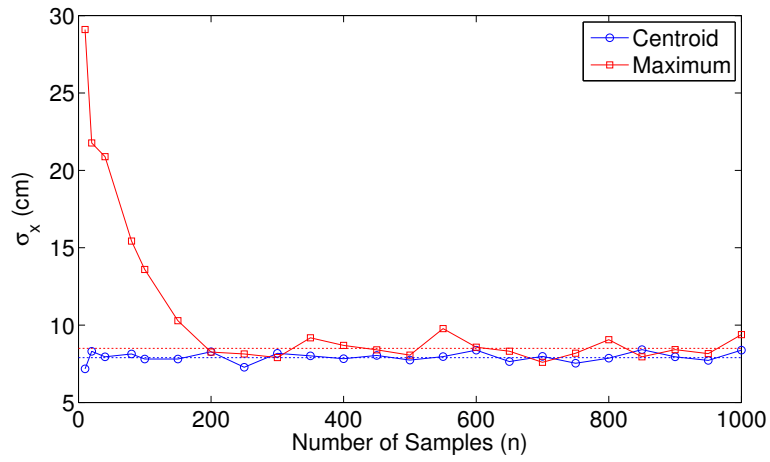
Figure 4.7: (a): Standard deviation of the offset between the true source direction and the centroid or maximum of the 3-D SBP image as a function of the number of reconstructed events, n. (b): (a) scaled by $\sqrt{n}$.

81

Here we will focus on the scenario where the arrays are placed far enough away from the source so that, for each individual array, the source is considered to be in the far field. This simplifying assumption allow the cones to be projected from the center of the detector array regardless of where the true initial interaction took place and also simplifies the implementation of the sensitivity.

4.3.1 MLEM Model Changes

Similar changes are made to the system model for the multiple-array geometry as were made for the single-array case, and (4.1) is still valid for multiple arrays.

The estimation of the sensitivity, however, is much simpler in this case than the single-array case. The strict definition of the sensitivity would require us to calculate the true system response in the 3-D space for all possible events in all of the arrays or estimate it by using the method proposed by Xu [5] (as was the case for the single-array case). However, since we are assuming that the source is in the far field for any given array, we can use the standard 2-D sensitivity of each array to estimate the sensitivity in the 3-D space. Being mindful of the spatial offset and orientation of the arrays, the 2-D (directional) sensitivity is projected into the entire 3-D space surrounding each array. To account for geometric attenuation, a $1/r^2$ scaling factor is included when the sensitivity is projected into the space where r is the distance between the origin of the array and the sample point in the image space. The projected sensitivities for each array are added together to form the final estimate of the total system sensitivity.

4.3.2 Available FOV

We perform a similar analysis that we did for the single array system in Section 4.2.2, but first we use a simple simulation to demonstrate the localization capability of the 3-D image reconstruction using two detector arrays. Figure 4.8 shows a
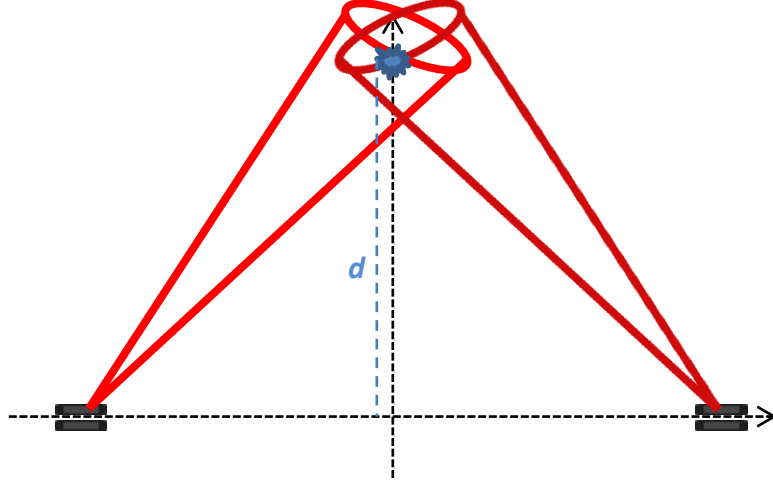
Figure 4.8: Diagram of the simulation of two detector arrays and a ^{137}Cs source.

diagram of the simulation geometry, and Figure 4.9 shows the reconstructed image using the standard MLEM algorithm for a distance d of 50 cm. The reconstruction successfully locates the source in 3D within 2 cm of the true simulated source position.

When using multiple detectors separated by a relatively large distance, the assumptions in Section 4.2.2 provide a more accurate model of the system than for a single detector array. Figure 4.10 shows the comparison between the FWHM of the source distance predicted by (4.2) and the measured FWHM from simulated data. Since the source is now in the far-field for both detector arrays, the reconstructed Compton cones from a given array all originate from essentially the same point, which is the key assumption made by (4.2), and the simulated results more closely match the predicted results. However, as the source gets closer to the detectors, near-field effects can cause the simulated FWHM to deviate slightly from the expected results.

For a single detector array, the distance estimate is the most limiting factor in localizing a point source, but other limitations arise for multiple-array geometries. A key example is when a source is placed directly between two detector arrays that are far apart. In this scenario, the reconstructed image would only reveal that the source lies between the two arrays. There is no way to determine where the source lies in between the two arrays without another measurement from a different angle.
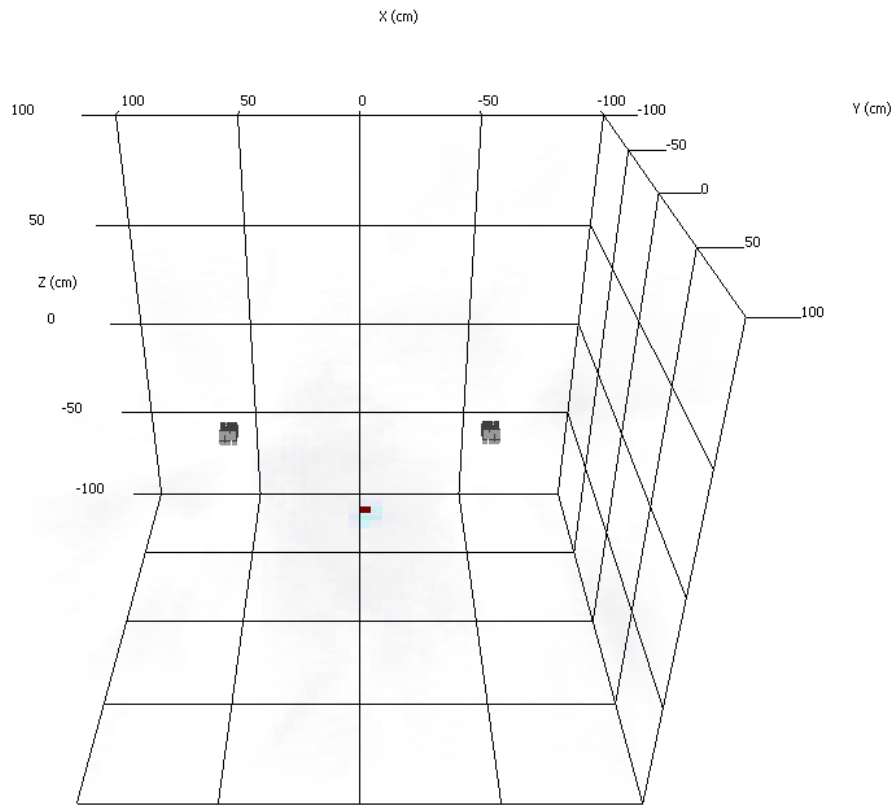
(a)
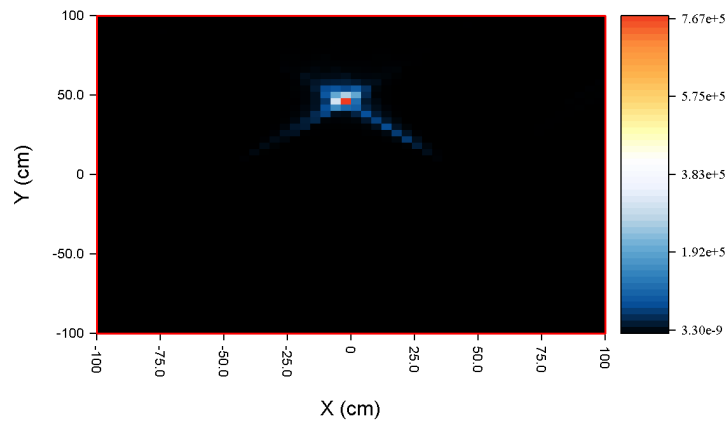
(b)

Figure 4.9: (a) 3-D rendering of the MLEM image reconstruction of a ^{137}Cs point source placed 50 cm perpendicular from the midpoint of two detector arrays. (b) X-Y slice (Z=0) of the reconstructed image shown in (a).
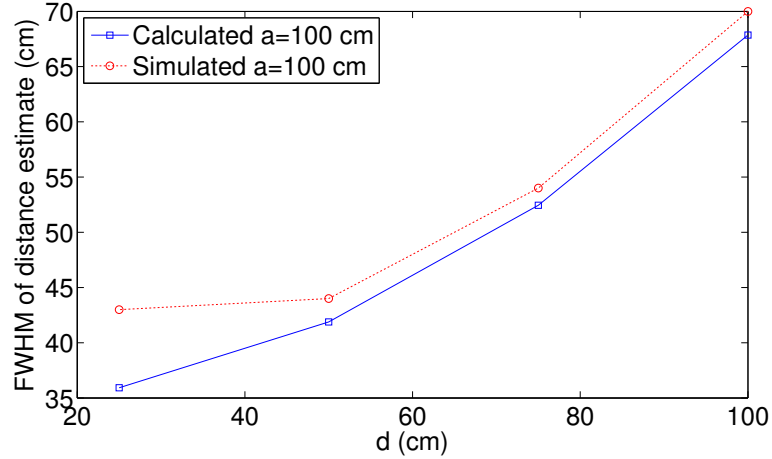
Figure 4.10: FWHM of the source distance estimate predicted by (4.2) and measured from reconstruction of simulated data.

Situations like this could be avoided if an array is mobile, which would allow it to get a better perspective and create more parallax. This type of reconstruction with a moving detector array is discussed in the following section.

4.4 Single Moving Array

In this section, we use the motion of a single detector array to create the parallax necessary to reconstruct sources in the 3-D space in the vicinity of the detector path. This type of system would allow an operator to map the radiation field in a room while he walks around without having to use crude $1/r^2$ techniques. Mihailescu et al. [46] performed a similar type of experiment with stop motion. In that work, they used a 3-D LIDAR system to map the room before performing the image reconstruction so that they could restrict the reconstruction space to solid objects present in the FOV. Although the system had the advantage of only voxellating the space that was likely to contain a source, it was complex and had to use first map the room with a LIDAR and then take measurements with their gamma-ray imaging system.

In this work, we simplify the problem by using a standard 3-D mesh for the reconstruction and projecting that mesh onto the standard 2-D mesh with an optical

overlay. The optical image allows the user to more easily interpret the image and quickly identify the source position in the room.

4.4.1 MLEM Model Changes

Several changes to the MLEM system model are necessary for this type of reconstruction. Just as with the other 3-D reconstructions, the system response function as well as the senstivity function need the addition of the $1/r^2$ term. However, since the array is in motion, the model must account for the time dependence of the reconstruction. In the following chapter, we derive the time-dependent list-mode log likelihood and a new MLEM algorithm for scenarios where the source is in motion. The most important change to the model is the fact that the sensitivity function s_j is now integrated over time. Of course for a stationary detector, this would be trivial, but since the array is moving for this work, the sensitivity needs to reflect the changes in the position and orientation of the array over time.

The system response function also has to account for any uncertainty in the detector position and orientation now that the system is in motion. To account for these uncertainties, we simply convolve the system response as caclulated in Section 3.3.2 with the uncertainty distribution in the spatial domain (*e.g.,* a 3-D Gaussian distribution). This simple post-process convolution is possible because, even though the motion introduces uncertainty in the absolute photon-interaction locations, there is is no added uncertainty in the relative positions of the gamma interactions. This distinction is important because if the relative positions added uncertainty, then we would have to account for it in the uncertainty of the cone axis. However, since the entire array moves as a single unit, the uncertainty caused by its motion manifests itself as a blur in the original stationary system response.

Any specific implementation would depend on the model of the uncertainty of the position and orientation. As seen in the next section, the position and orientation

estimation used for this work does not have a well-characterized uncertainty, but the uncertainty is relatively small compared to the those already present in the system model. Therefore, we assume this uncertainty is negligible and it is not included in the model for this work.

4.4.2 Position Estimation

Since the detector array is now in motion, some way to estimate the position and orientation of the array is necessary in order to correctly reconstruct the Compton cones in the 3-D space. One may immediately think of GPS as an obvious choice to estimate the position and orientation of the array. However, GPS is a poor choice because it does not work indoors, and typical standalone GPS units only have a precision on the order of several meters [47]. These limitations make GPS unsuitable for some of the applications of interest for our detector array (like mapping the radiation field in a building).

Instead of using GPS for this proof-of-principle work, we used Robosense, a system designed for mobile robotics in the late 1990's, to estimate the position and orientation of the detector array. The Robosense unit uses a laser range finder in combination with a spinning mirror that measures the distance and relative angle of stationary reflective targets (see Fig. 4.11) in the field-of-view. With this information, the unit can map the room and estimate its own lateral position and heading. This information is timestamped and recorded at a rate of roughly 10 Hz. For this work, the array is sitting on a cart with wheels so the elevation of the detector is assumed to be constant. Also, the only rotation considered is rotation about the vertical, z, axis which corresponds to the direction that the Robosense unit records. Other rotations about the x or y axes are ignored since it is laying flat on the surface of the cart. If the detector was truly handheld, we would need to consider the other rotational information.

<div align="center">(a)   (b)</div>

Figure 4.11: (a) Robosense unit and (b) reflective target for the Robosense unit.

The Robosense unit was able to yield fairly accurate position and heading estimates. some simple experiments were performed to test the accuracy of the position estimation and found that the standard deviation of the position resolution was about 5 cm in both the x and y directions, and the standard deviation of the heading was about 0.5°. However, there was some offset present in both measurements. The heading had a constant offset of about 5° that was easily corrected, but the position estimate had an offset that was position dependent on the order of 5-15 cm in either direction. Since the overall uncertainty in both position and heading is mostly negligible in comparison to the angular uncertainty of the Compton cone reconstruction, we simplified the problem and neglected them in the model.

4.4.3 2-D Projection

Reconstructing the 3-D Compton image in a standard rectangular mesh is simple and convenient. However, it is does not give the operator a way to quickly correlate the source position in the reconstructed image to an object in physical space. When displaying 2-D Compton images, we can overlay the radiation image over a 4π optical image as seen in Figure 4.12. This optical image allows the user to easily locate the

(a)

(b)
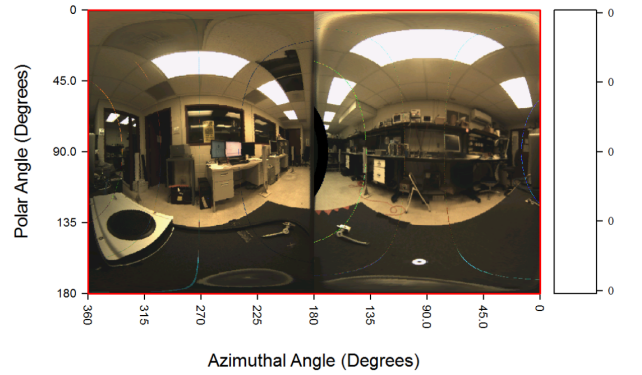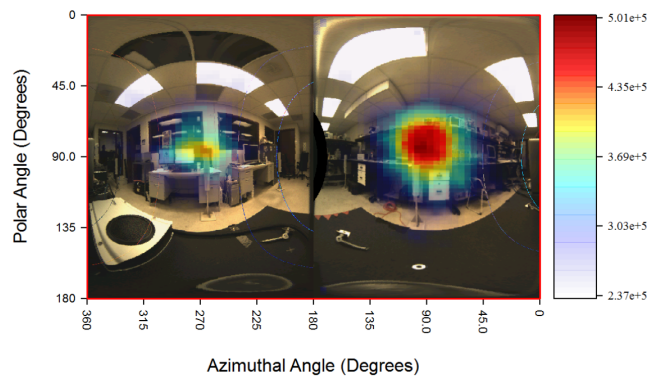
Figure 4.12: (a) 4π optical image of the space surrounding the detector array. (b) Gamma-ray image overlaid on the on the optical image.

source direction in the real world.

Overlaying the radiation image on the 4π optical image is straight forward when reconstructing on the standard 2-D directional mesh. To utilize the optical image for 3-D reconstructions, we must calculate the projection of the 3-D radiation image on the 2D sphere as a function of position and orientation of the detector system which is changing continuously.

The projection calculation is relatively simple. First, a simple coordinate transform is performed to represent the position of each 3-D voxel of the reconstruction in the reference frame of the detector array. Next, we calculate the polar and azimuthal angles of the relative position of each 3-D voxel and add the intensity of that voxel to

the nearest pixel on the 2-D imaging sphere surrounding the array. After the entire 3-D space has been added to the pixels in the 2-D imaging sphere, the intensity of each pixel is normalized to the number of voxels that contributed to it. This normalization ensures that the intensity of each pixel is not related to the distance to a boundary in that direction, but it also tends to reduce the contrast of the hotspots in the image.

Even though this view only allows the user to see the direction of the source as in the standard 2D case, this method shows the image from many different angles as the detector moves. This way the reconstruction can continuously integrate as the detector is in motion, and the operator can observe how the hotspot tracks with the objects in the room while in motion without having to take stop-motion measurements with long dwell times.

4.4.4 Experiment

Figure 4.13 shows a simple schematic of the experiment we performed to test this reconstruction. A 30 μCi ^{137}Cs source and a 30 μCi ^{22}Na source were placed on opposite sides of the room at the same elevation as the detector array. The array itself was positioned on top of a cart which was free to move and pivot in the room. The motion experiment lasted approximately 4.5 minutes and about 4,000 imageable photopeak counts were recorded.

The motion path taken by the detector system during the experiment is also shown in Figure 4.13. An initial sweep of the room was taken, straight forward and back, in order to get a good assessment of the overall radiation field. Once the initial sweep was finished, paths closer to the two sources were then taken in order to get a better estimate of the source location with more perspective. After the close up view of the sources, a final sweep forward and back is taken and the measurement ended.

Figures 4.14 and 4.15 show the cumulative simple back projection results using
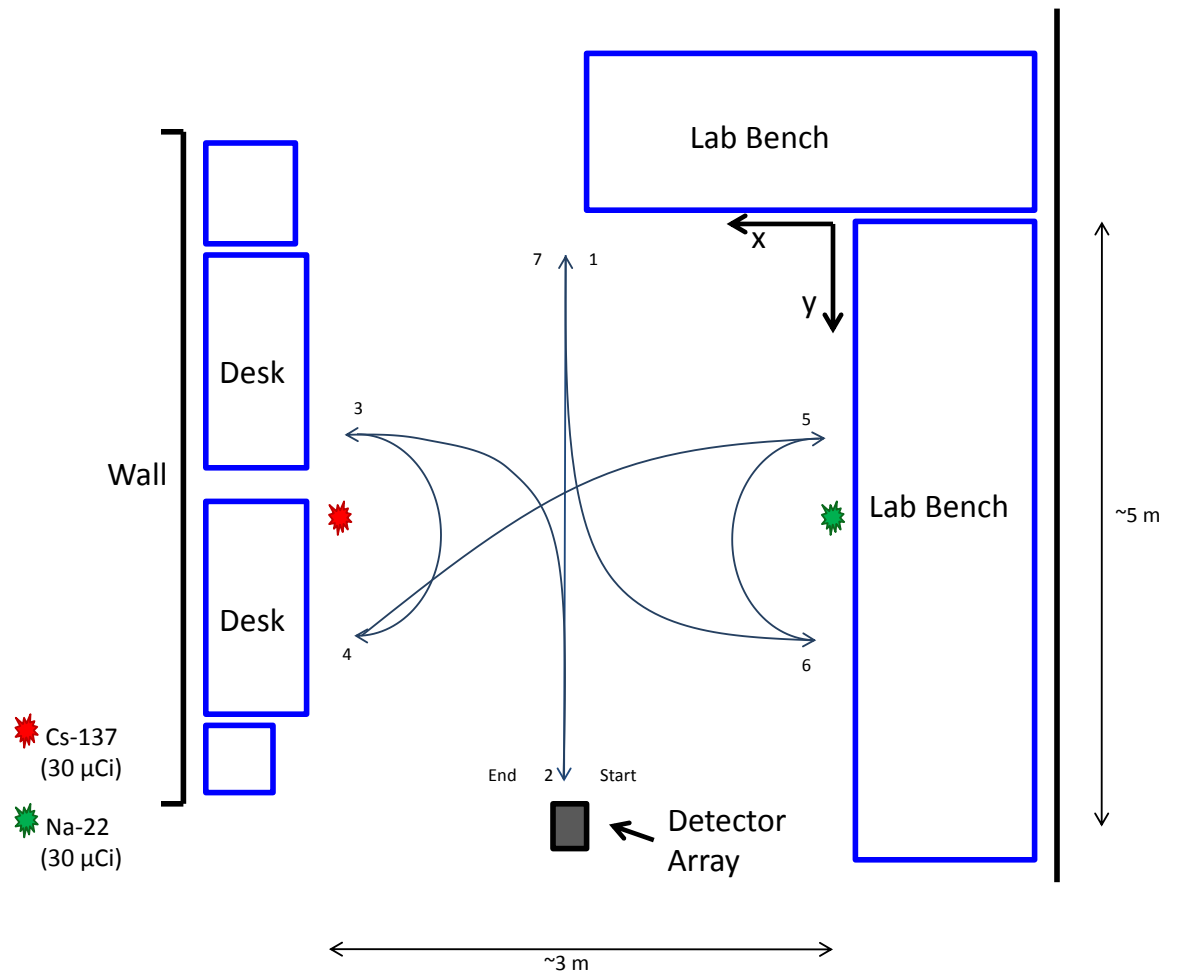
Figure 4.13: Simple schematic of the experiment and description of the detector motion. The motion started and stopped at the same point and moved through the field following the numbered points.

photopeak counts (511 keV, 662 keV, and 1274 keV) as a function of time which coincide with points 2, 4, and 6 and the start and end points shown in Figure 4.13. For each time there are two images shown. First the 2D projection of the 3D mesh is shown overlaid on the optical image, and second is the full 3D image shown with the current position and orientation of the detector array. Figures 4.14(a) and 4.14(b) show the initial position and optical image before the measurement began. Figure 4.14(a) also shows the true location of the sources in the optical image.

As time progresses, the location of the sources becomes immediately obvious. Note that the reconstructed source positions track the true position of the sources as seen in the optical image. Some offset is noticeable in some images, and two reasons account for this inaccuracy. The refresh rate of the optical cameras is sometimes as slow as several seconds. This would cause the optical image to not reflect the true image for relatively large segments of time. Also, the optical cameras are not placed exactly where the detectors are located. As such, there is a noticeable amount of mismatch between the optical image and the reconstructed radiation image whenever the objects in the optical image are within about a one meter radius of the detector. Despite these subtle inaccuracies, the locations of the sources are readily identified in the optical and 3D images. The ^{22}Na source, which emits approximately three times as many photons as the ^{137}Cs source, appears more intense as expected.

For comparison, Figure 4.16 shows the results from the MLEM reconstruction of the ^{137}Cs photopeak events using the standard system model. As in the 2-D reconstructions, the MLEM reconstruction improves the spatial resolution of the reconstructed image, and the reconstructed hotspot in this case is accurate to within 10 cm. The MLEM reconstruction of the ^{22}Na photopeak counts produced similar results as seen in Figure 4.17.

4.5 Discussion

In this chapter, we have shown the capability to localize sources in 3-D for a variety of detector configurations, and we were able to successfully demonstrate the use of a detector in continuous motion to reconstruct the source distribution in 3-D space.

These 3-D image reconstruction algorithms can be applied to a variety of problems. Near-field imaging with a single detector array has obvious applications in the medical field [48]. Multiple stationary detector arrays could be useful for surveillance

(a)

(b)

(c)

(d)
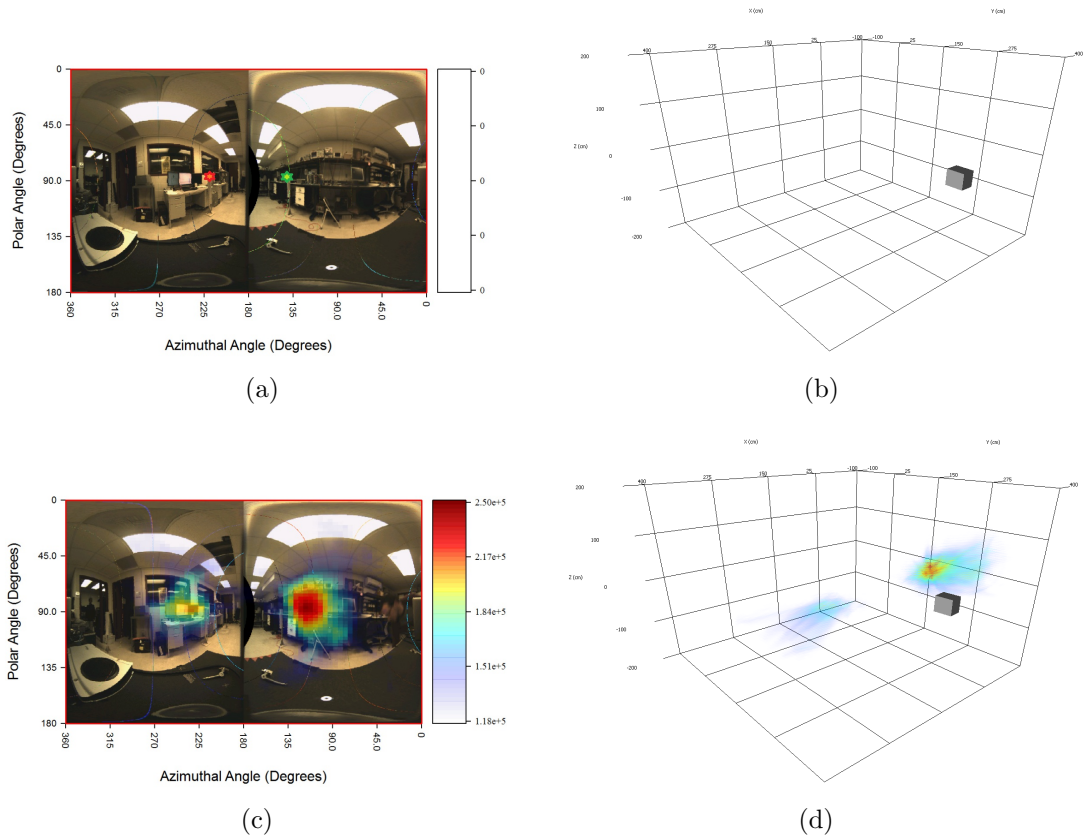
Figure 4.14: Cumulative SBP image at different times. Images on the left are 2D projections of the 3D imaging mesh overlaid on the optical image. Images on the right show the 3D images as well as the location and orientation of the detector array system. (a) Optical image before measurement (with sources marked) and (b) the initial detector position. (c) The image and (d) detector position after the initial forward and back motion (point 2 on Figure 4.13).

(a)

(b)

(c)

(d)

(e)

(f)
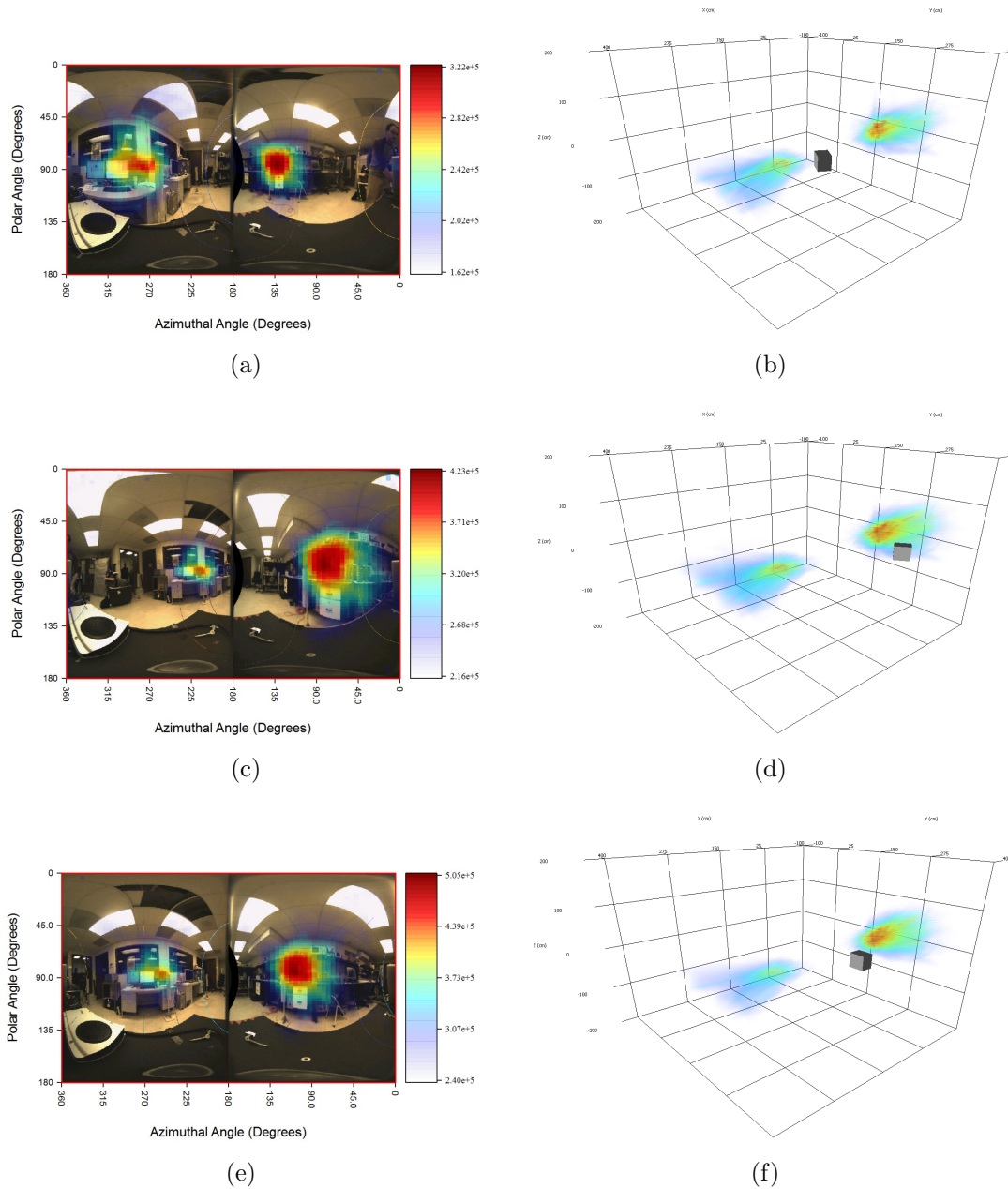
Figure 4.15: Cumulative SBP image at several different times. Images on the left are 2D projections of the 3D imaging mesh overlaid on the optical image. Images on the right show the 3D images as well as the location and orientation of the detector array system. (a) The image and (b) detector position after the close sweep of the ^{137}Cs source (point 4 on Figure 4.13) (c) The image and (d) detector position after the close sweep of the ^{22}Na source (point 6 on Figure 4.13). (e) The final image and (f) detector position.

94
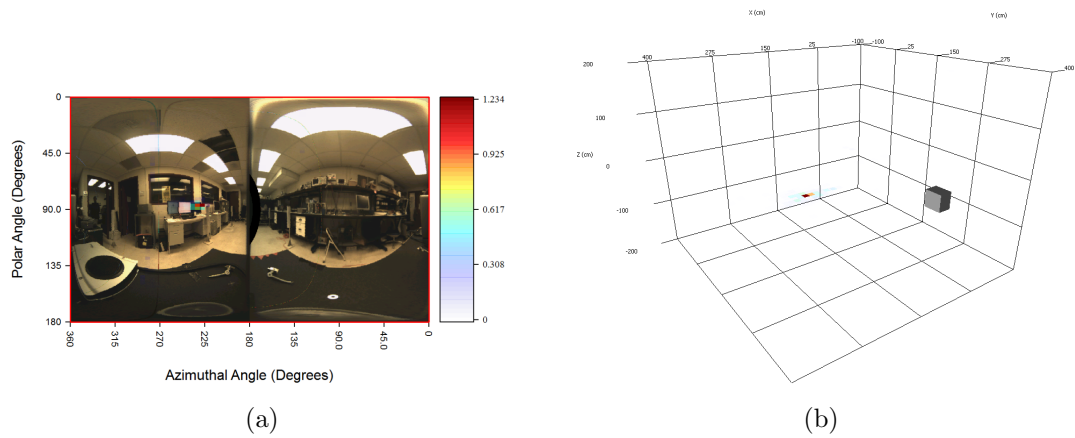
Figure 4.16: Final MLEM reconstruction results using the standard system response for the ^{137}Cs photopeak events.
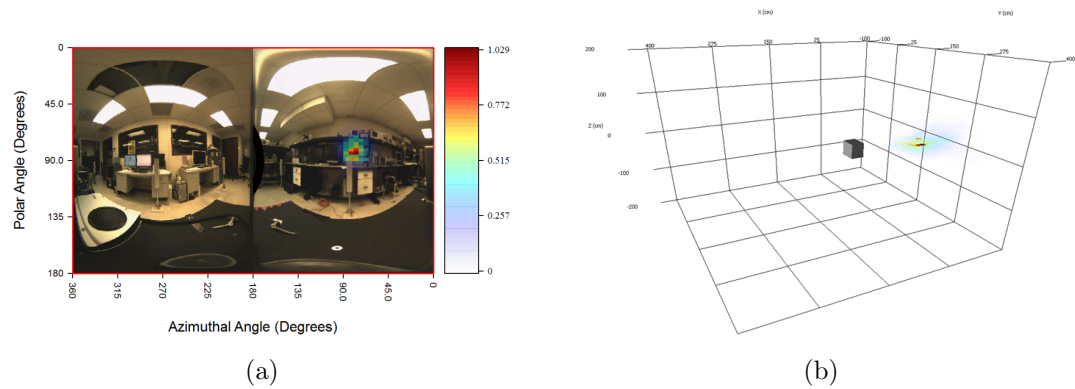
Figure 4.17: Final MLEM reconstruction results using the standard system response for the ^{22}Na photopeak events.

in nuclear non-proliferation or for secondary inspection in homeland security, and moving detectors are well suited for search scenarios.

Of course, in the real world many times the deployment of radiation detectors is more complicated. One can imagine a variety of detector types including counters, spectrometers, and imaging spectrometers all of which can be moving or stationary. This complicated scenario of many different detectors composes a "sensor network" [49]. The work presented in this chapter lays the ground for the ability to incorporate all of the information from this sensor network to reconstruct a single source distribution in 3D.

However, these complicated and likely large 3-D imaging domains introduce computational challenges, especially when the detector is allowed to move freely. As the imaging space expands, simple imaging meshes become large (especially if fine spatial resolution is required), and the addition of the energy and/or time domains would exacerbate the problem. These challenges must to be addressed if the 3-D algorithms need to be robust enough to handle arbitrarily large spaces.

CHAPTER V

Image Reconstruction with Known Source Motion

The previous chapter culminated with an experiment using a moving detector to reconstruct the 3-D locations of point sources. When the detector is in motion, presumably it is aware of its own location, and that motion allows the 3-D source position to be reconstructed. However, in some cases the sources are in motion instead of the detector, and for these cases the detector system has no inherent knowledge of the position of the source as a function of time. Thus, the reconstructed image of the moving point source will result in a convolution of the source path and the point spread function of the reconstruction method. However, if the profile of the source motion is known, the reconstruction can use that information to compensate the reconstructed image for that motion.

This chapter will briefly discuss some current methods of motion compensation followed by two different ways to compensate images for known source motion. Finally a new technique for motion compensation is introduced which allows for simultaneous reconstruction of many objects with independent motion paths without blur and cross contamination.

5.1 Description of Previous Work

In the medical field there has been active research on source-motion compensation for many years. One obvious way to account for motion is to explicitly include the time domain in the reconstruction, *e.g.,* 4-D PET [50]. Often this approach uses a standard pixel mesh in the spatial domain and then uses B-splines to parameterize the time domain [51]. The resulting reconstruction represents the time-dependent activity distribution revealing physiological processes that would otherwise be unobservable with a single 3-D image. Variants of this approach have been studied, including using Fourier basis functions [52] and using list-mode data to improve temporal resolution [53], as well as many others.

Other techniques simply attempt to remove blur from an image by using motion compensation instead of reconstructing a full time-dependent image. Bloomfield *et al.* tracked patient movement by an optical tracking system and corrected the lines of response for this motion before image reconstruction [54]. Later Rahmim *et al.* expanded on the work to include the motion compensation in the reconstruction model itself [55], which among other things, accounted for the change in sensitivity of the system as the patient moves. This improvement was an important step to improve the accuracy of the system model.

In other applications, such as homeland security, imaging and/or detection of gamma-ray sources contained in moving objects is desired. For example, one can imagine placing detectors near border crossings or other choke points to detect illicit nuclear material carried in vehicles or by pedestrians. In this case, motion compensation would be required to faithfully reconstruct the intensity distributions around these (moving) objects. In certain situations, there may also be stationary sources of gamma rays residing in what we refer to as the backdrop that are also of interest for imaging and detection. This approach requires that the motion of the object(s) of interest are known through some other means of estimation such as video tracking

systems [56–58].

5.2 Motion Compensation Techniques

There are different ways that one may choose to account for source motion. Of course if that motion is small compared to the resolution of the imaging system, attempting to compensate for that motion is most likely not a top priority. However, when the motion is non-negligible, choosing to not compensate for the motion will result in excess blur in the image. In an extreme case, as seen in Figure 5.1(a), where a ^{137}Cs point source traveled all the way around the detector array with constant angular velocity, the reconstructed SBP image shows a smear through the equator corresponding to the source motion. In this case, the motion was slow enough that it is obvious in the reconstructed image, but if the source moved at a higher velocity (or if the source was weaker), the source motion would be less obvious as seen in Figure 5.1(b).

5.2.1 Simple Motion Compensation

One way to account for the source motion is to continuously rotate the reference frame of the detector to keep the current source position in the center of the field of view. Simply put, the algorithm translates the position of the reconstructed ring in the imaging domain to account for the motion of the source as seen in Figure 5.2. This type of compensation works well when there is a single source present. When multiple sources are present, this simple algorithm has no way to keep all of the sources in focus (unless they all have the same motion relative to the detector). Figure 5.3 shows the reconstructed image of the same data used for the reconstructed image in Figure 5.1 except that there is a stationary ^{22}Na source located above the equator. In this case a choice must be made as to which source will be in focus. If no motion compensation is performed (Figure 5.3(a)) the ^{22}Na source is in focus. If the reconstruction accounts
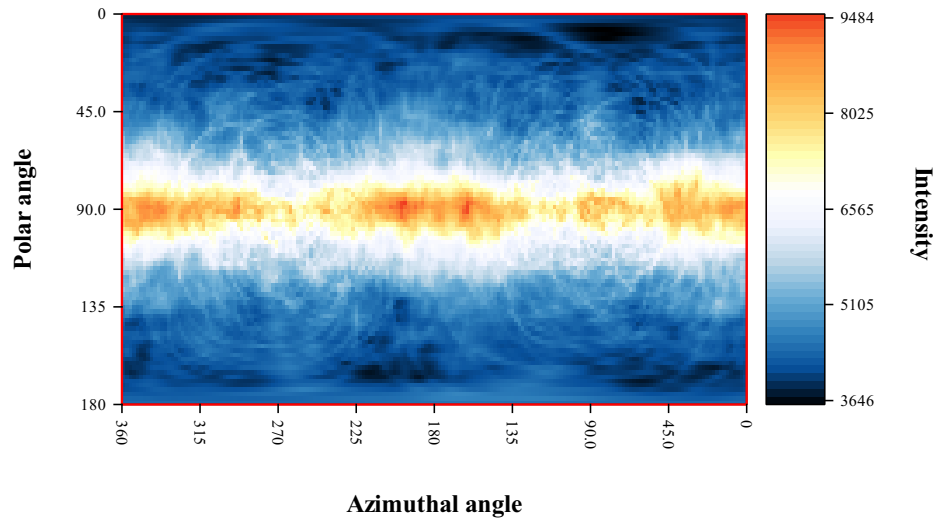
(a)
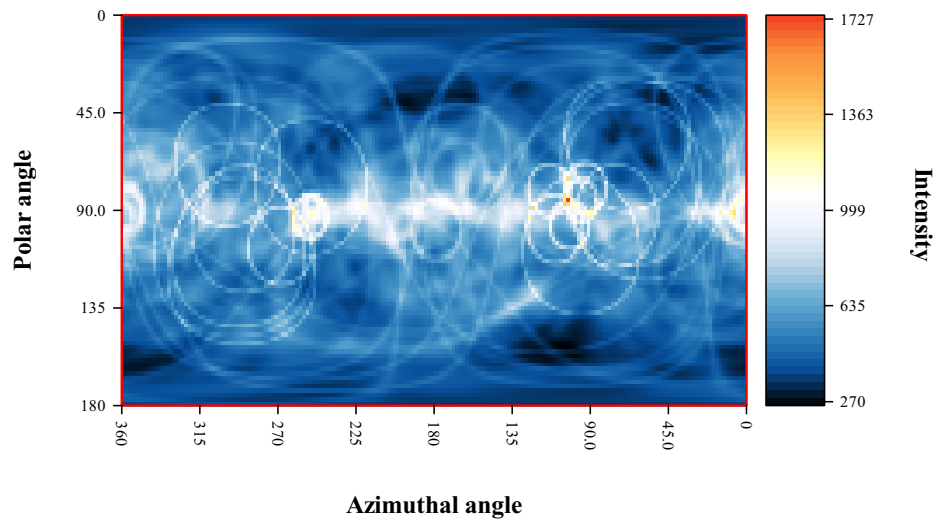
(b)

Figure 5.1: (a) SBP image of a ^{137}Cs point source that moved 360° around the detector array at constant angular velocity. (b) SBP image of a ^{137}Cs point source that moved 360° around the detector array at constant angular velocity 10 times faster than in (a).
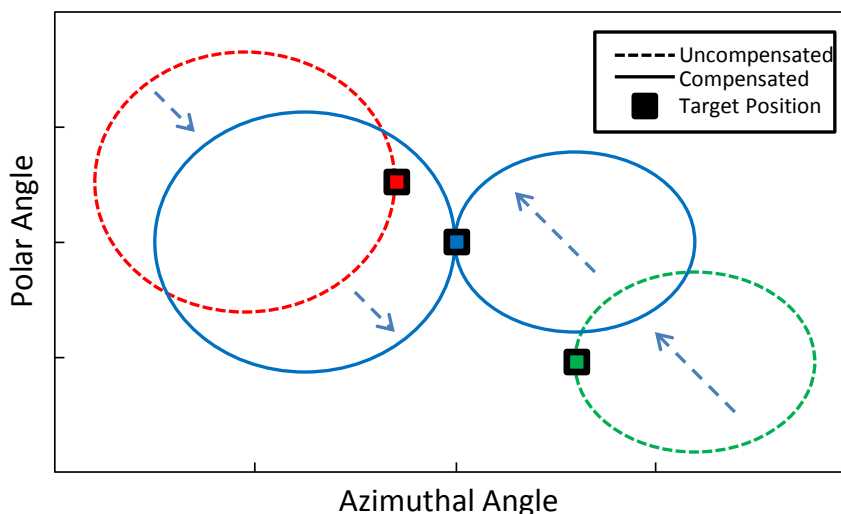
Figure 5.2: Simple example of Compton rings translated to account for the source motion.

for the ^{137}Cs motion (Figure 5.3(b)) the ^{22}Na source is smeared.

5.2.2 Target Spatial Binning

To avoid the necessity of choosing just one object to keep in focus, a new way of binning the image must be created. The idea is to create a new *target* image domain that will be a subset of the standard imaging domain. By creating this new domain, the target space can be compensated for source motion while leaving the standard imaging domain, or *backdrop* space, unaltered.

Figure 5.4 demonstrates how this compensation is performed. When a Compton ring is reconstructed, it is added to the backdrop without any motion compensation because it is possible that the photon came from a source in the stationary backdrop. If part of the Compton ring happens to intersect a direction corresponding to the target domain at the time that the interaction occurred, the fraction of the ring that overlaps the target is added to the target space. In this way, if the target space accurately tracks the source direction, the Compton rings resulting from full-energy
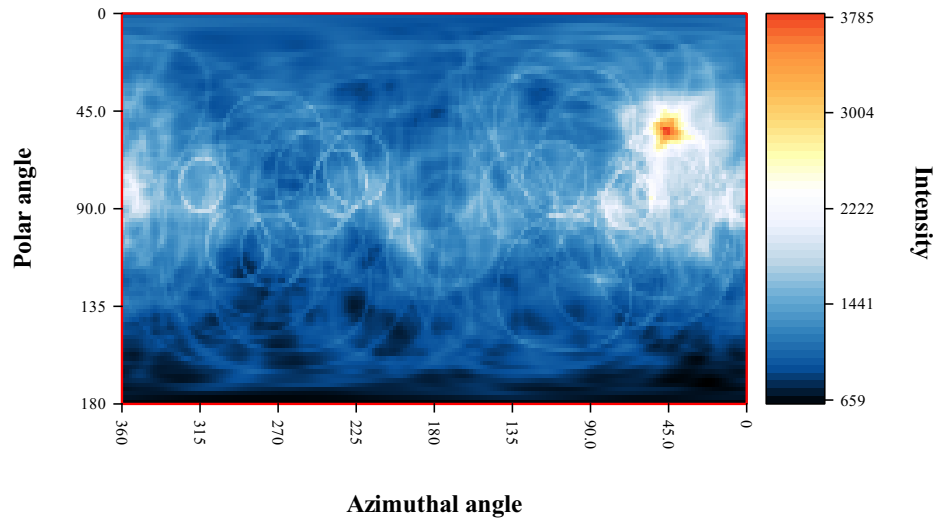
(a)
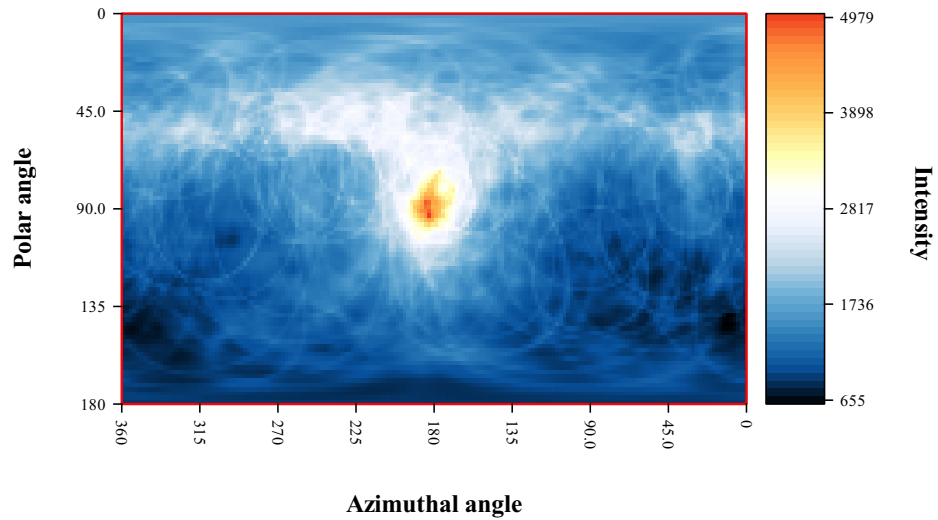
(b)

Figure 5.3: (a) Uncompensated SBP image of a ^{137}Cs point source that moved 360°
around the detector array at constant angular velocity at the equator and
a stationary ^{22}Na point source. (b) SBP of the data used for the image in
(a) using simple motion compensation. Note that the ^{137}Cs source now
appears to be stationary at the center of the image, but the ^{22}Na source
is smeared in the azimuthal direction.
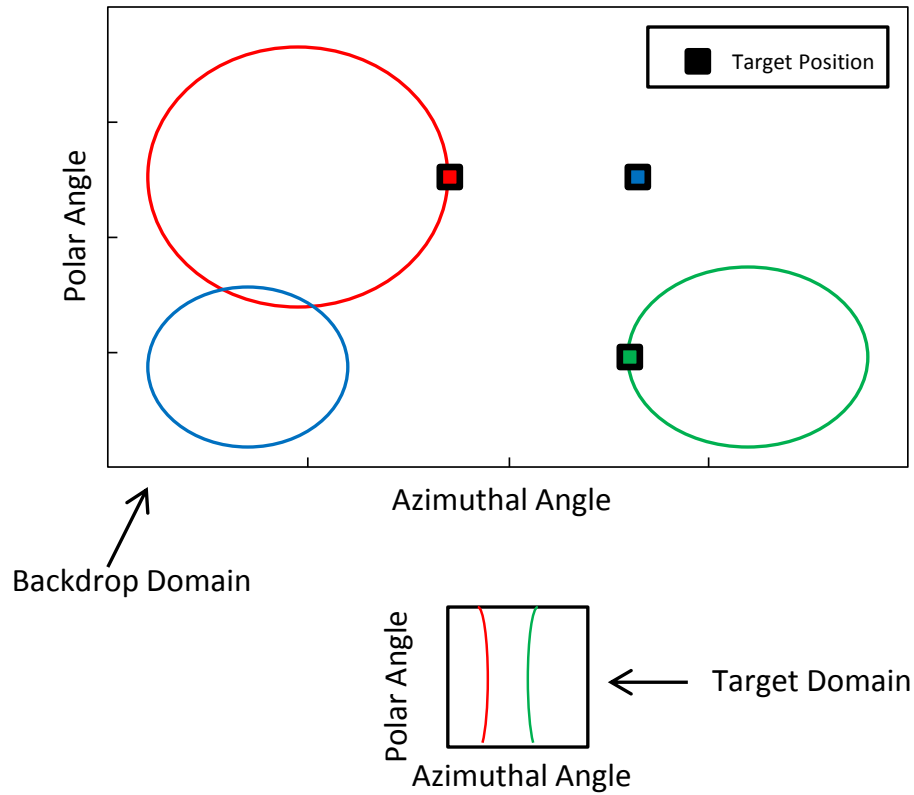
Figure 5.4: Simple example of how the proposed target spatial binning accounts for source motion without affecting the distribution in the backdrop space.

deposition events from the moving source will be added together as if the source were stationary without causing any stationary sources to be blurred in the backdrop space.

5.3 Derivation

This algorithm can be implemented easily for the SBP reconstruction where the rings are simply added together in the imaging space, and the imaging pixels are completely independent of each other. However, for the MLEM based reconstructions, some care must be taken to ensure that the log-likelihood correctly accounts for the time-dependency of the new imaging voxels in the target space since the standard list-mode log-likelihood does not explicitly incorporate time-dependencies. Also, the

expectation maximization algorithm must be re-derived for the new likelihood expression.

5.3.1 Model

The terms used for this derivation include those found in Section 3.3 modified to include time as well as other new terms.

5.3.1.1 Data Model

The attribute vector $\boldsymbol{v}$ recorded by the detector system still represents the recorded energies and 3-D positions of the photon interactions, but now time is recorded in addition to the $\boldsymbol{v}$ vector. Thus, the list-mode acquisition for a fixed scan duration τ_d records M attributes $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_M$ at corresponding times $t_1, \ldots, t_M$.

5.3.1.2 System Model

The system sensitivity function is now also dependent on time since part of the imaging mesh will change as a function of time

$$s(\vec{r}, E, t) = \mathsf{P}[D \mid \vec{r}, E, t]. \tag{5.1}$$

Now this term is interpreted as the probability that a gamma-ray emission from spatial position $\vec{r}$ at energy E *and* time t is detected (D) by the system.

Of course the other model component is the system response function,

$$f_m(\vec{r}, E, t_m) = \mathsf{p}(\boldsymbol{v}_m \mid \vec{r}, E, t_m, D_m)s(\vec{r}, E, t_m) \tag{5.2}$$

which describes the probability of recording $\boldsymbol{v}_m$ for emissions of energy E originating from spatial position $\vec{r}$ for a specific event m. The specific expression of this term is shown in (3.25)-(3.27). In this case, the time dependence is only used to correct the

spatial position $\vec{r}$ for the target motion as described below.

5.3.1.3   Target Object Model

For a static (stationary) scene, the usual goal is to reconstruct the radioactive source intensity distribution $\lambda(\vec{r}, E)$ in some spatial-energy domain Ω from the measured list-mode data, and the existing list-mode approach [35, 36, 39] is appropriate. The new method which we derive here extends this formulation to the case where the scene consists of one or more target objects that are moving relative to a stationary backdrop, *i.e.*, the overall source intensity distribution is a function of time: $\lambda(\vec{r}, E, t)$. We shall consider here the case of "known" motion, *i.e.*, the motion is estimated separately, for example by a video tracking system. Note that the time dependence is only present in the motion of the target through space. The intensities of the sources are assumed to be constant over the scan time in this work.

We model the intensity distribution as consisting of an unknown stationary backdrop intensity distribution $\lambda_0(\vec{r}, E)$ and a set of K target objects that may be moving. The unknown intensity distribution of the kth object at time $t = 0$ is denoted

$$\lambda_k(\vec{r}, E) \equiv \lambda_k(\vec{r}, E, 0) \tag{5.3}$$

for $k = 1 \ldots K$. For simplicity of presentation, we assume that the target-object motion can be modeled adequately by a spatial translation. Other forms of motion, such as rotation, could be accommodated as well. The spatial shift of the kth target object at time t is denoted $\vec{c_k}(t)$, where $\vec{c} \in \mathbb{R}^3$ for a 3-D imaging problem or $\vec{c} \in \mathbb{R}^2$ for the 2-D case. Under this assumption, the intensity distribution of the kth target object at time t is modeled as

$$\lambda_k(\vec{r}, E, t) = \lambda_k(\vec{r} - \vec{c_k}(t), E). \tag{5.4}$$

Also for simplicity, this model ignores occlusions and the absorption of gamma photons by target objects that move in front of one another. Finally, the overall time-varying intensity distribution of the scene is modeled as the superposition of the stationary backdrop $\lambda_0(\vec{r}, E)$ and the moving target-object contributions:

$$
\begin{aligned}
\lambda(\vec{r}, E, t) &= \lambda_0(\vec{r}, E) + \sum_{k=1}^{K} \lambda_k(\vec{r}, E, t) \\
&= \lambda_0(\vec{r}, E) + \sum_{k=1}^{K} \lambda_k(\vec{r} - \vec{c}_k(t), E). \quad (5.5)
\end{aligned}
$$

The goal is to reconstruct the intensity distributions of the backdrop and the target objects, $i.e.,$ $\{\lambda_0(\vec{r}, E), \lambda_1(\vec{r}, E), \ldots, \lambda_K(\vec{r}, E)\}$, from the list-mode data, assuming known motion $\vec{c}_k(t)$ for $k = 1, \ldots, K$.

To facilitate numerical implementation, the intensity distributions are parameterized using a finite-series model [59]:

$$
\lambda_0(\vec{r}, E) = \sum_{j=1}^{N_0} \lambda_{0j} b_{0j}(\vec{r}, E), \quad (5.6)
$$

where N_0 is the total number of basis functions in the backdrop, λ_{0j} denotes the unknown intensity of the jth basis function and b_{0j} denotes the basis function (typically a 2-D region of space in the spherical coordinates or a voxel in the 3-D coordinates, covering an energy interval). Similarly the moving target objects are parameterized (at time $t = 0$) as:

$$
\lambda_k(\vec{r}, E) = \sum_{j=1}^{N_k} \lambda_{kj} b_{kj}(\vec{r}, E), \quad (5.7)
$$

where the number of basis functions N_k used to represent the kth target object may differ between target objects of different sizes. If the kth moving target object is treated as a point source, then N_k reduces to the number of energy bins. The overall
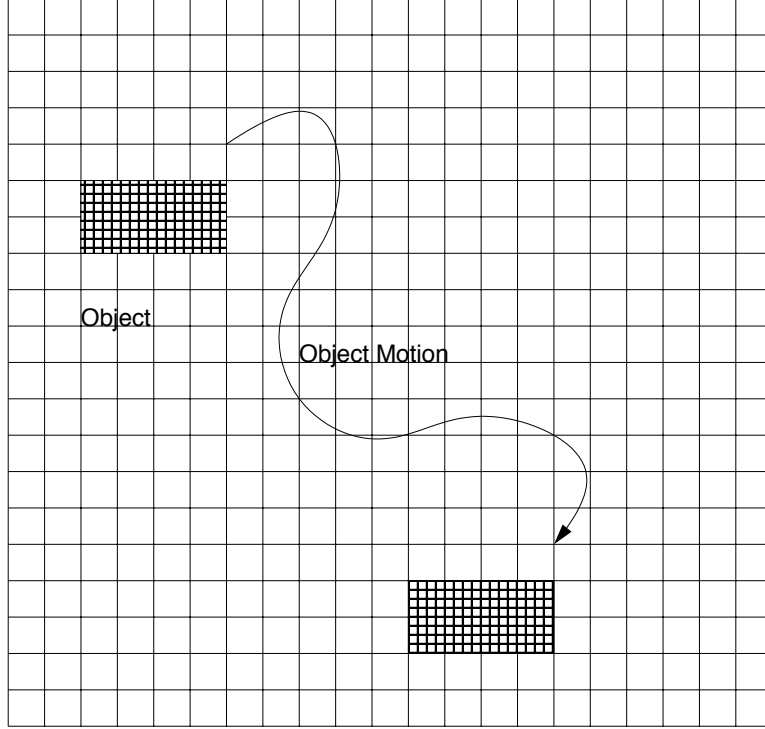
Figure 5.5: Illustration of target object model with a pixellated stationary backdrop and one pixellated target object that moves during the imaging process.

time-varying intensity distribution is thus parameterized as:

$$\lambda(\vec{r}, E, t) = \sum_{k=0}^{K} \sum_{j=1}^{N_k} \lambda_{kj} b_{kj}(\vec{r} - \vec{c_k}(t), E) \tag{5.8}$$

where the backdrop mesh has been included over the target sum for concision and $\vec{c_0}(t) \equiv 0$. Typically, the basis functions are unitless and the coefficients λ_{kj} have units of "emissions per unit time per unit solid angle (or volume) per unit energy." With this parameterization, the goal is to estimate the coefficients of the backdrop $\{\lambda_{0j}\}$ and of the targets $\{\lambda_{kj}, k = 1, \ldots, K\}$ from the list-mode data.

Figure 5.5 illustrates the model. Note that if the kth target object does not move, then this formulation could be over-parameterized because the intensity within the support of that target object could be modeled both by the target object pixel λ_{kj} and a corresponding backdrop pixel $\lambda_{0j'}$, in the case that $b_{kj}(\vec{r})$ and $b_{0j'}(\vec{r})$ overlap.

Therefore, one should include only moving target objects in this formulation otherwise the solution will not be unique. (Stationary sources will be reconstructed as part of the backdrop.)

5.3.2   Time-Dependent List-Mode Log Likelihood

With these tools, the general form of the list-mode log likelihood for time-varying emission distribution $\lambda(\vec{r}, E, t)$ is derived following the form of a similar derivation found in [60]. Assume that the detection system is set to record photon interaction events for a preset time. In this case, the number of counts the system records will be a Poisson random variable. We also assume that the detector records the time t_m of the mth recorded event as well as a set of attributes $\boldsymbol{v}_m$ describing the event including the number of interactions as well as the locations and energies of each of those interactions. The log likelihood associated with these observations is

$$
\begin{aligned}
\mathsf{L}(\lambda) = \ & \mathsf{log}\bigg(\mathsf{p}\big((\boldsymbol{v}_1, t_1 \mid D_1), (\boldsymbol{v}_2, t_2 \mid D_2), \ldots \\
& \ldots, (\boldsymbol{v}_n, t_n \mid D_n) \mid M = n; \lambda\big) \\
& \cdot \mathsf{P}\big[M = n, D_1, D_2, \ldots, D_n; \lambda\big]\bigg).
\end{aligned}
\tag{5.9}
$$

where n is the total number of events recorded, D_m denotes that the mth event is detected, and λ is the intensity distribution in space and energy from (5.8). This expression describes the log of the joint probability of recording the given attributes, given that they were detected and that exactly n events were detected, multiplied by the probability that n events were detected. The following term, which describes the expected count rate as a function of time, is defined to simplify future expressions

$$
\lambda_s(t) = \iint_\Omega s(\vec{r}, E, t)\lambda(\vec{r}, E, t)\,\mathrm{d}\vec{r}\mathrm{d}E.
\tag{5.10}
$$

The last term in the log likelihood, which follows a Poisson distribution, describes the probability of recording exactly $M = n$ counts during the scan time. Thus,

$$P[M = n, D_1, D_2, \ldots, D_n; \lambda] = \frac{e^{-\bar{M}_\lambda} (\bar{M}_\lambda)^n}{n!} \qquad (5.11)$$

where $\bar{M}_\lambda$ is the expected number of recorded counts for source disribution λ. Since $\lambda_s(t)$ is the instantaneous expected count rate, the expected total number of photons recorded over the scan time τ_d is:

$$\bar{M}_\lambda = \int_0^{\tau_d} \lambda_s(t) \, \mathrm{d}t. \qquad (5.12)$$

The first part of the list-mode log likelihood can be rewritten using the chain rule as:

$$p\big((\boldsymbol{v}_1, t_1 \mid D_1), \ldots, (\boldsymbol{v}_n, t_n \mid D_n) \mid M = n; \lambda\big) =$$
$$p(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \mid t_1, \ldots, t_n, D_1, \ldots, D_n, M = n; \lambda)$$
$$\cdot\, p(t_1, \ldots, t_n \mid D_1, \ldots, D_n, M = n; \lambda). \qquad (5.13)$$

Note that the recorded time t_m of each event cannot be included in the attribute vector $\boldsymbol{v}_m$ because the times follow a specific order and thus cannot be considered independent parameters. If we assume the attribute vectors are conditionally independent of each other given the event times (meaning we ignore dead time and pile-up), the first term on the right hand side of (5.13) can be futher simplified to:

$$p(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \mid t_1, \ldots, t_n, D_1, \ldots, D_n, M = n; \lambda) =$$
$$\prod_{m=1}^{n} p(\boldsymbol{v}_m \mid t_m, D_m; \lambda). \qquad (5.14)$$

The second term in (5.13) describes the conditional distribution of the ordered

arrival times. For a Poisson random process Ross [61, pg. 37,53] shows that

$$\mathsf{p}(t_1, t_2, \ldots, t_n, D_1, D_2, \ldots, D_n \mid M = n; \lambda) =$$

$$\begin{cases} n! \prod_{m=1}^{n} q(t_m; \lambda) & 0 < t_1 < t_2 < \cdots < t_n < \tau_d \\ \\ 0 & \text{otherwise} \end{cases} \tag{5.15}$$

where

$$q(t_m; \lambda) =$$

$$\begin{cases} \frac{\lambda_s(t_m)}{\bar{M}_\lambda} & 0 < t_1 < t_2 < \cdots < t_n < \tau_d \\ \\ 0 & \text{otherwise.} \end{cases} \tag{5.16}$$

Substituting (5.11)-(5.16) into (5.9) yields the following simplified expression for the list-mode log-likelihood for time-varying sources:

$$\mathsf{L}(\lambda) = \sum_{m=1}^{n} \log\left(\mathsf{p}(\boldsymbol{v}_m \mid t_m, D_m; \lambda) \, \lambda_s(t_m) \right) - \bar{M}_\lambda. \tag{5.17}$$

Using the logarithm of the likelihood in (5.9) is a key factor in simplifying the likelihood expression. It allows the product to be reduced to a sum (which will be important for the EM derivation in the next section), and it allowed for easier simplification of the Poisson distribution in (5.11).

To analyze $\mathsf{p}(\boldsymbol{v}_m \mid t_m, D_m; \lambda)$, we use total probability:

$$\mathsf{p}(\boldsymbol{v} \mid t, D; \lambda) =$$

$$\iint_\Omega \mathsf{p}(\boldsymbol{v} \mid \vec{r}, E, t, D; \lambda) \mathsf{p}(\vec{r}, E \mid t, D; \lambda) \, \mathrm{d}\vec{r} \, \mathrm{d}E. \tag{5.18}$$

110

The first term in the integral is simply

$$\mathsf{p}(\boldsymbol{v} \mid \vec{r}, E, t, D; \lambda) = \mathsf{p}(\boldsymbol{v} \mid \vec{r}, E, t, D) \qquad (5.19)$$

which is the distribution of recorded attributes expected from a source at location $\vec{r}$ and energy E recorded at time t. It is independent of the overall intensity distribution and is a key term in the system model (see (5.2)).

The second term in the integral is the probability density function (pdf) for the emission of a photon at position $\vec{r}$ and energy E. This term is related to λ but is a bit more complicated as a result of the conditioning on D, because the photon is only *recorded* if the photon is detected. Particularly, the conditional distribution of the origin of the incident photons is:

$$\mathsf{p}(\vec{r}, E \mid t, D; \lambda) = \frac{\mathsf{P}[D \mid \vec{r}, E, t; \lambda]\mathsf{p}(\vec{r}, E \mid t; \lambda)}{\mathsf{P}[D \mid t; \lambda]}. \qquad (5.20)$$

The first term in the numerator is simply the conditional sensitivity of the system at time t which is independent of λ, and the second term in the numerator is the pdf of the source intensity distribution $\mathsf{p}_t(\vec{r}, E)$, and is directly proportional to $\lambda(\vec{r}, E, t)$ since the decay constant of the radioactive isotopes is assumed to be long compared to the measurement time. Finally, the denominator is the total probability of detection at time t which is calculated by integrating the sensitivity multiplied by $\mathsf{p}_t(\vec{r}, E)$. Thus, (5.20) reduces to:

$$\begin{aligned}
\mathsf{p}(\vec{r}, E \mid t, D; \lambda) &= \frac{s(\vec{r}, E, t)\mathsf{p}_t(\vec{r}, E)}{\underset{\Omega}{\iint} s(\vec{r}\,', E')\mathsf{p}_t(\vec{r}\,', E') \, \mathrm{d}\vec{r}\,' \, \mathrm{d}E'} \\
&= \frac{s(\vec{r}, E, t)\lambda(\vec{r}, E, t)}{\lambda_s(t)}.
\end{aligned} \qquad (5.21)$$

Substituting(5.19) and (5.21) into (5.18) yields:

$$p(\boldsymbol{v} \mid t, D; \lambda) =$$
$$\frac{1}{\lambda_s(t)} \iint_{\Omega} p(\boldsymbol{v} \mid \vec{r}, E, t, D) s(\vec{r}, E, t) \lambda(\vec{r}, E, t) \, \mathrm{d}\vec{r} \, \mathrm{d}E. \tag{5.22}$$

After substituting (5.22) into (5.17), the time-dependent log-likelihood expression is:

$$L(\lambda) = \sum_{m=1}^{n} \log \left(\iint_{\Omega} p(\boldsymbol{v}_m \mid \vec{r}, E, t_m, D_m) s(\vec{r}, E, t_m) \right.$$
$$\left. \cdot \lambda(\vec{r}, E, t_m) \, \mathrm{d}\vec{r} \, \mathrm{d}E \right) - \bar{M}_\lambda \tag{5.23}$$

which using (5.2) simplifies to

$$L(\lambda) = \sum_{m=1}^{n} \log \left(\iint_{\Omega} f_m(\vec{r}, E, t_m) \lambda(\vec{r}, E, t_m) \, \mathrm{d}\vec{r} \, \mathrm{d}E \right) - \bar{M}_\lambda \tag{5.24}$$

5.3.3 EM Algorithm

Starting with (5.24) and using optimization transfer techniques [7,60,62], we derive the EM update equation. The goal is to be able to write the maximization in a simple way separable in λ_{kj}. First, write the log likelihood in terms of the parameterized

intensity distribution (5.8):

$$\mathsf{L}(\lambda) =$$

$$\sum_{m=1}^{n} \log\bigg(\iint\limits_{\Omega} \mathsf{p}(\boldsymbol{v}_m \mid \vec{r}, E, t_m, D_m) s(\vec{r}, E, t_m)$$

$$\cdot \sum_{k=0}^{K} \sum_{j=1}^{N_k} \lambda_{kj} b_{kj}(\vec{r} - \vec{c}_k(t_m), E) \, \mathrm{d}\vec{r}\mathrm{d}E \bigg)$$

$$- \int_{0}^{\tau} \iint\limits_{\Omega} s(\vec{r}, E, t) \sum_{k=0}^{K} \sum_{j=1}^{N_k} \lambda_{kj} b_{kj}(\vec{r} - \vec{c}_k(t), E) \, \mathrm{d}\vec{r}\mathrm{d}E \, \mathrm{d}t. \qquad (5.25)$$

Now, rearrange and make simplifying substitutions:

$$\mathsf{L}(\lambda) = \sum_{m=1}^{n} \log\bigg[\sum_{k=0}^{K} \sum_{j=1}^{N_k} \lambda_{kj} f_{kjm} \bigg] - \sum_{k=0}^{K} \sum_{j=1}^{N_k} \lambda_{kj} s_{kj} \qquad (5.26)$$

where,

$$f_{kjm} = \iint\limits_{\Omega} \mathsf{p}(\boldsymbol{v}_m \mid \vec{r}, E, t_m, D_m) s(\vec{r}, E, t_m) b_{kj}(\vec{r} - \vec{c}_k(t_m), E) \, \mathrm{d}\vec{r}\mathrm{d}E \qquad (5.27)$$

$$s_{kj} = \int_{0}^{\tau} \iint\limits_{\Omega} s(\vec{r}, E, t) b_{kj}(\vec{r} - \vec{c}_k(t), E) \, \mathrm{d}\vec{r}\mathrm{d}E \, \mathrm{d}t. \qquad (5.28)$$

Introduce a new scalar term, β_{kjm}, and rewrite (5.26) :

$$\mathsf{L}(\lambda) =$$

$$\sum_{m=1}^{n} \log\bigg[\sum_{k=0}^{K} \sum_{j=1}^{N_k} \frac{\beta_{kjm} \lambda_{kj} f_{kjm}}{\beta_{kjm}} \bigg] - \sum_{k=0}^{K} \sum_{j=1}^{N_k} \lambda_{kj} s_{kj} \qquad (5.29)$$

where

$$\beta_{kjm} > 0 \quad \forall \, k, j, m \qquad \text{and} \qquad \sum_{k=0}^{K} \sum_{j=1}^{N_k} \beta_{kjm} = 1 \quad \forall \, m.$$

Note that $\log(x)$ is a concave function, *i.e.*:

$$\log\left(\sum_i a_i x_i\right) \geq \sum_i a_i \log(x_i) \tag{5.30}$$

under the following conditions:

$$\forall x_i, \quad a_i > 0, \quad \sum_i a_i = 1 \quad \text{and} \quad \sum_i a_i x_i \neq 0. \tag{5.31}$$

Taking advantage of this property, we can write a surrogate function that is both separable in λ_{kj} and no greater than the original log likelihood:

$$\mathsf{L}(\lambda) \geq \Phi(\lambda; \beta_{kjm}) = \sum_{k=0}^{K} \sum_{j=1}^{N_k} \phi_{kj}(\lambda_{kj}; \beta_{kjm}) \tag{5.32}$$

where

$$\phi_{kj}(\lambda_{kj}; \beta_{kjm}) = \sum_{m=1}^{n} \beta_{kjm} \log\left[\frac{\lambda_{kj} f_{kjm}}{\beta_{kjm}}\right] - \lambda_{kj} s_{kj}. \tag{5.33}$$

The idea here is at iteration i to initially set β_{kjm}^{i} such that $\Phi(\lambda^i; \beta_{kjm}^i) = \mathsf{L}(\lambda^i)$ and maximize the surrogate function with respect to λ, which is easier than maximizing the original likelihood function. Then calculate β_{kjm}^{i+1} using λ^{i+1}. Repeat this process until the maximum in the surrogate is equal to (or "near") the maximum of the log-likelihood. To implement this algorithm, one needs to know how to update β_{kjm} and maximize $\phi_{kj}(\lambda_{kj}; \beta_{kjm}^i)$ with respect to λ_{kj}.

First set $\Phi(\lambda^i; \beta_{kjm}) = \mathsf{L}(\lambda^i)$ and solve for β_{kjm}^i:

$$\beta_{kjm}^i = \frac{\lambda_{kj}^i f_{kjm}}{\sum_{k'=0}^{K} \sum_{j'=1}^{N_{k'}} \lambda_{k'j'}^i f_{k'j'm}}. \tag{5.34}$$

Next, maximize $\phi_{kj}(\lambda_{kj}; \beta^i_{kjm})$ to find λ^{i+1}_{kj}:

$$\frac{\partial}{\partial \lambda_{kj}} \phi_{kj}(\lambda_{kj}; \beta^i_{kjm}) = 0 = \sum_{m=1}^{n} \frac{\beta^i_{kjm}}{\lambda_{kj}} - s_{kj}$$

$$\lambda^{i+1}_{kj} = \frac{1}{s_{kj}} \sum_{m=1}^{n} \beta^i_{kjm}. \tag{5.35}$$

Combine (5.34) and (5.35) to obtain the list-mode EM algorithm for moving sources:

$$\lambda^{i+1}_{kj} = \frac{\lambda^i_{kj}}{s_{kj}} \sum_{m=1}^{n} \frac{f_{kjm}}{\sum_{k'=0}^{K} \sum_{j'=1}^{N_{k'}} \lambda^i_{k'j'} f_{k'j'm}} \tag{5.36}$$

A key difference between this result and the update equations derived by Parra and Barrett [36] is the time integral of the system sensitivity. Due to motion of the object and/or the system itself, we compute (5.28) by Riemann sum over a list of discretized target and detector positions and rotations. An example where this distinction is important is when a target gets very close to the system during the measurement time. If relatively few counts are recorded during that period, the instantaneous sensitivity $s(\vec{r}, E)$ would be large for only a small fraction of events, but the overall sensitivity s_{kj} would be large and that high overall sensitivity would suppress the reconstruction from estimating a source that passes that direction.

5.4   Performance

To test the performance of the newly proposed algorithm, we performed several experiments using the standard 18-detector array described in Section 2.4. Here we discuss two of those experiments. The first experiment demonstrates the ability to compensate for source motion without disturbing the reconstruction of stationary sources, and the second shows the capability of reconstructing multiple moving sources without introducing cross contamination between them.
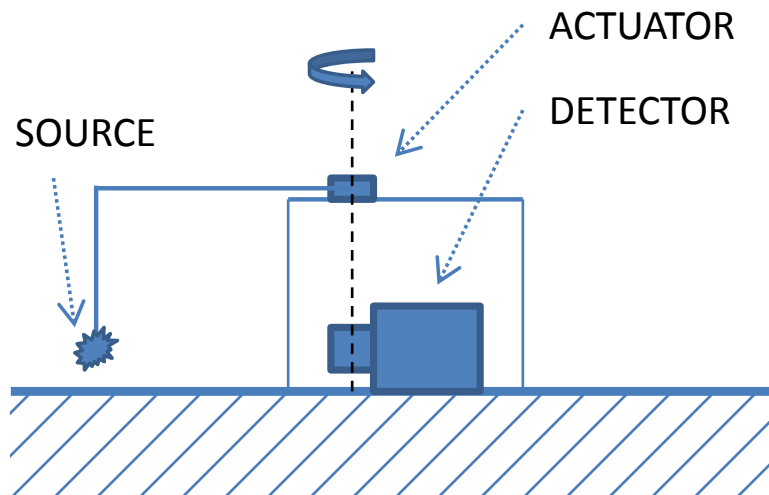
Figure 5.6: A computer controlled actuator arm which is mounted to a table is used to control the position of the source in motion. The detector system is depicted as two rectangles. The smaller rectangle contains the actual CdZnTe crystals, and the larger body contains the supporting electronics including high and low voltage power supplies and the data readout. The array of detectors is positioned directly underneath the actuator pivot point, and the source is placed at the bottom of the actuator arm, in-plane with the detectors.

5.4.1 Target-Position Estimation

To demonstrate the presented algorithm, a method of recording source position as a function of time was required. Thus, the apparatus shown in Figure 5.6 was assembled to rotate a source around the 18-detector array. A computer controlled actuator, attached to the top of a table, was set to rotate with a constant angular velocity to known angular displacements. An 'L'-shaped aluminum arm was attached to this actuator so that the bottom-most tip of the arm was in the plane of the array system positioned below the table. The center of the detector head was placed directly below the actuator pivot point so that the resulting motion at the bottom-most tip of the actuator arm would revolve completely around the center of the detector array.
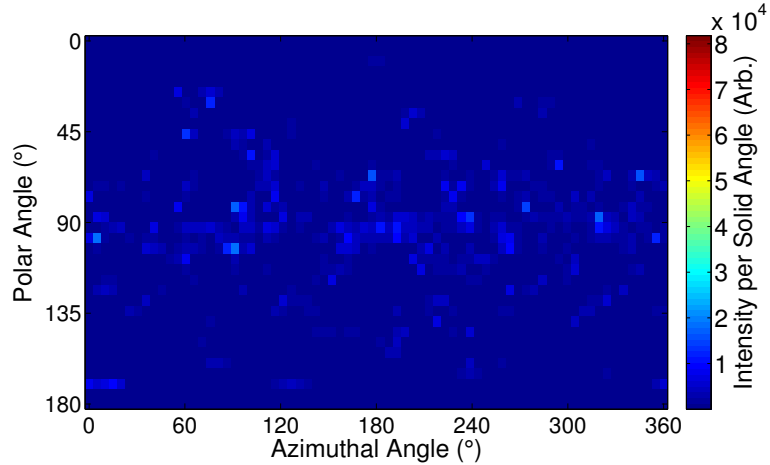
Figure 5.7: Results for the standard MLEM spatial reconstruction for the ^{137}Cs energy slice. The estimated intensity is smeared through the equator since no motion compensation was applied and the source was in constant motion.

5.4.2 One Moving and One Stationary Source

In the first experiment, a 122-μCi ^{137}Cs source was rotated 360° at a constant velocity around the detector system at a radius of 1.2 m during the course of a 46-minute measurement in the counterclockwise direction (when viewed from the top). Also, a [22]Na source was placed in a stationary position in the backdrop 1.0 m from the center of the detector. Because of computational constraints, a smaller set of events was used for the following reconstructions. To use data from the entire data-collection time, the first event of every 50 events was used for each reconstruction of this data set which resulted in about 4000 imageable counts in the energy range of 300 keV to 1300 keV.

An MLEM reconstruction using the standard system response (which reconstructs only the spatial domain) is performed to compare with the new algorithm presented in this work. No motion compensation is performed, and energy windows are used so that only photopeak counts are reconstructed.

Figures 5.7 and 5.8 show the reconstructed images of the standard MLEM recon-
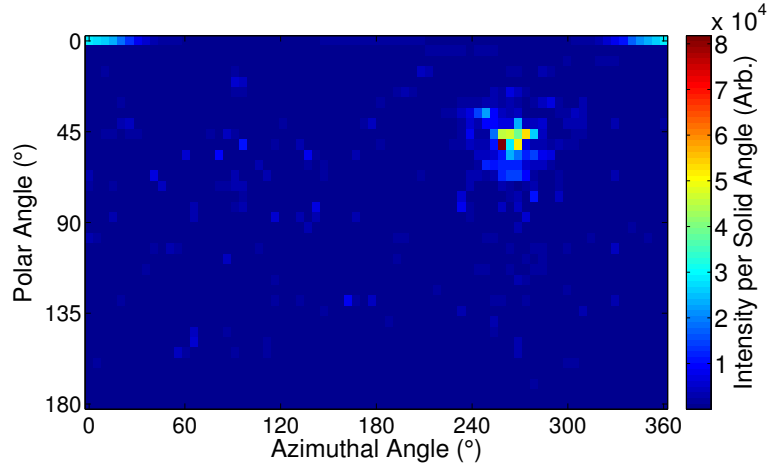
Figure 5.8:
Results for the standard MLEM spatial reconstruction for the ^{22}Na energy slice. A single localized hotspot is visible corresponding to the location of the stationary ^{22}Na source.

struction after 20 EM iterations using a 36×72 pixel backdrop imaging mesh (over 4π). Figure 5.8 shows that the spatial distribution for the ^{22}Na photopeak energy range is deconvolved correctly as expected since the source was stationary. However, the deconvolved spatial distribution for the ^{137}Cs photopeak window, seen in Figure 5.7 does not show a clear hotspot since the source was in constant motion and that motion was not included in the model.

The newly proposed method reconstructed the same data using a similar backdrop mesh and a 9×9 pixel target-object mesh spanning $40°$ in the polar and azimuthal directions. The energy domain consisted of 250 evenly spaced energy bins over a range of 300 - 1300 keV. Few imagable events fall outside this energy range, so the energy dimension could be limited to this set of energies.

Figures 5.9-5.11 show the reconstructed images after 20 EM iterations of the proposed model-based algorithm. The desired results in this situation would estimate a single hotspot in the stationary backdrop with an energy distribution corresponding to a ^{22}Na spectrum and a hotspot at 662 keV in the target object space, which tracked the ^{137}Cs source motion. The actual reconstruction for the backdrop space, seen in
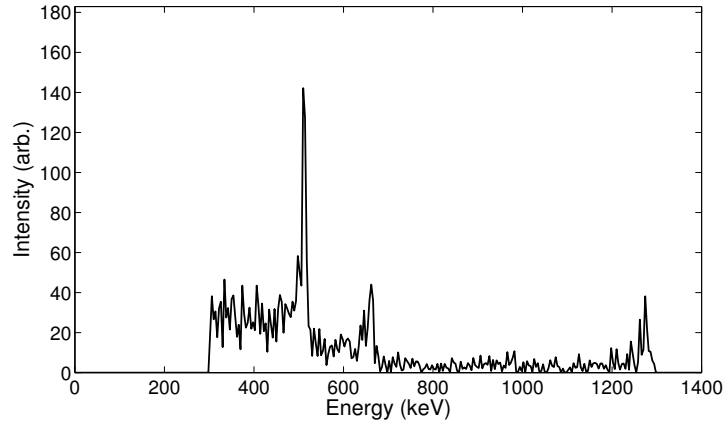
(a)
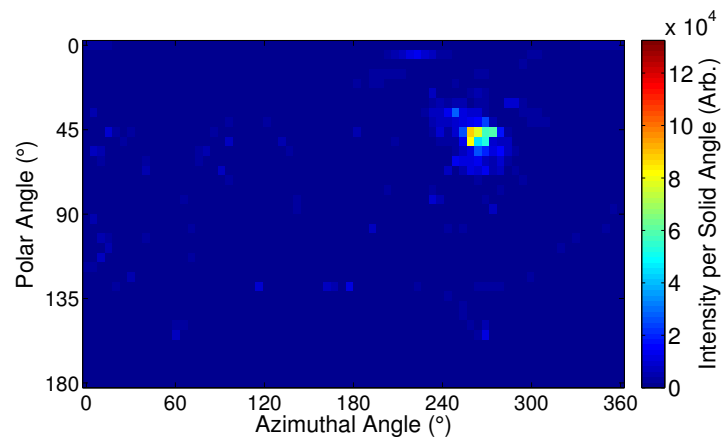
(b)
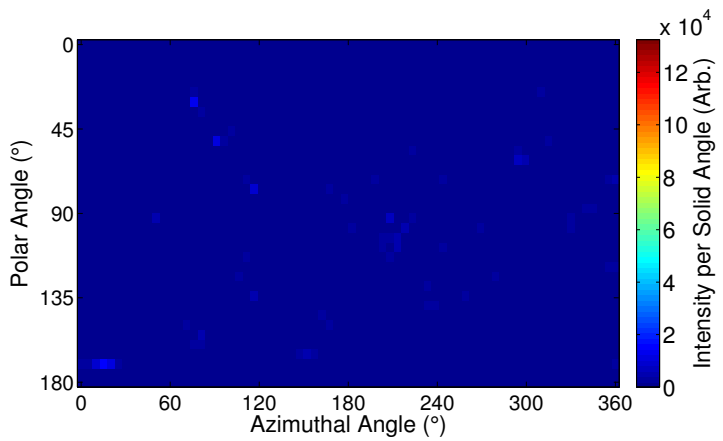
(c)

Figure 5.9: Results for the backdrop space after 20 iterations of the proposed motion-compensated EM algorithm for a moving ^{137}Cs source and a stationary ^{22}Na source: (a) reconstructed incident energy spectrum for the entire backdrop space; (b) reconstructed spatial distribution for the ^{22}Na energy slice; (c) reconstructed spatial distribution for the ^{137}Cs energy slice.
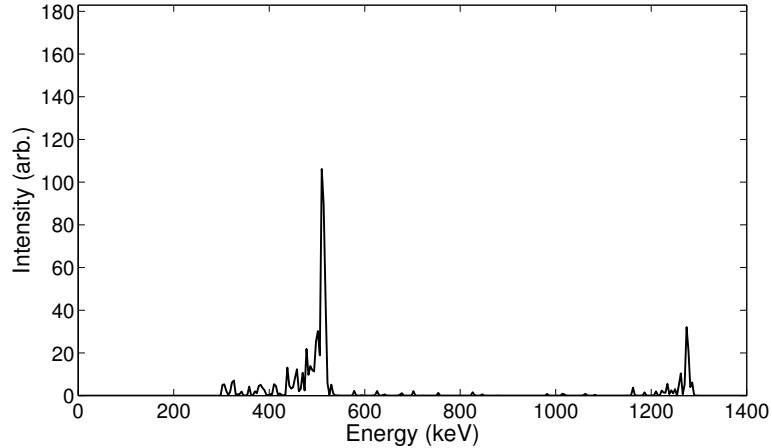
Figure 5.10: The reconstructed incident energy spectrum corresponding to just the hotspot direction in Figure 5.9(b).

Figure 5.9, is similar to the desired result. Figure 5.9(b) shows the single hotspot at the ^{22}Na energies, and even though there is again some ^{137}Cs contamination in the backdrop seen in Figure 5.9(c), it does not form a single hotspot but is disbursed over the entire 4π space. Figure 5.9(a) shows the reconstructed spectrum for the entire backdrop space, but it has a significant amount of ^{137}Cs contamination. Figure 5.10 shows the spectrum just in the direction of the stationary ^{22}Na source, which is a well-deconvolved ^{22}Na spectrum as expected. Figure 5.11 shows the reconstruction results for the target object space. The reconstructed spectrum shown in Figure 5.11(a) is a well deconvolved ^{137}Cs spectrum free of ^{22}Na contamination. Figure 5.11(b) shows the hotspot at the ^{137}Cs energy, and Figure 5.11(c) shows that there is no cross talk at the ^{22}Na energies.

5.4.3 Two Sources Moving in Opposite Directions

In the second experiment, the same ^{137}Cs source was rotated 360° around the detector system similar to the first experiment, and then the previously stationary ^{22}Na source was rotated 360° around the detector in the opposite direction in a similar fashion. The two datasets were then combined and reconstructed as if they occurred si-

120

(a)

Intensity per Solid Angle (Arb.)
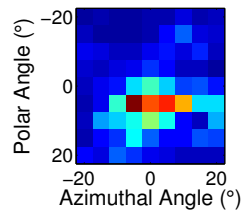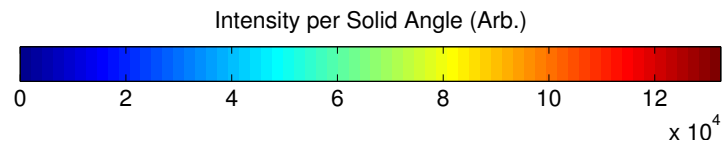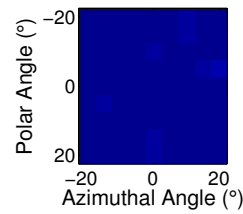
(b)   (c)

Figure 5.11: Results for the target-object space after 20 iterations of the proposed motion-compensated EM algorithm for a moving ^{137}Cs source and a stationary ^{22}Na source: (a) reconstructed incident energy spectrum for this target; (b) reconstructed spatial distribution for the ^{137}Cs energy slice; (c) reconstructed spatial distribution for the ^{22}Na energy slice.

multaneously with the sources crossing paths at 180° in the azimuthal direction. This combined dataset, consisting of about 4500 imageable events in a similar energy range (after downsampling), was reconstructed using similar imaging and energy meshes as the previous experiment except that now there are two target-object meshes tracking the two moving sources. The standard non-motion-compensated MLEM image for this dataset looks very similar to that of the first reconstruction: a smear of intensity through the equator, and without tracking information, the two sources would appear to simply overlap.

The results for the proposed model-based algorithm after 20 EM iterations are shown in Figures 5.12 and 5.13. The desired results for this reconstruction would be zero (or some small amount of background) in the backdrop space, and a single hotspot at the ^{137}Cs energy in the first target object which tracked the ^{137}Cs source, and a single hotspot at the ^{22}Na energies for the second target object which tracked the ^{22}Na source. The backdrop results seen in Figure 5.12 show roughly the desired results. Some ^{137}Cs and ^{22}Na intensity is incorrectly estimated in the backdrop, but this contamination is small and dispersed in all directions. Figure 5.13 shows the results for the two target objects. Here it is obvious that the first target object has a single hotspot at 662 keV and a zero distribution at the ^{22}Na energy window. Also, the second target object has a hotspot at the ^{22}Na energies and a flat distribution at the ^{137}Cs energy as desired. Finally, the estimated incident spectra for the two target objects show well deconvolved ^{137}Cs and ^{22}Na energy spectra.

5.5 Discussion

This chapter introduced a new imaging algorithm that can successfully reconstruct motion compensated images for multiple sources with independent motion profiles without cross contamination between the moving sources and any sources present in the stationary backdrop image. The reconstruction works well for point sources when
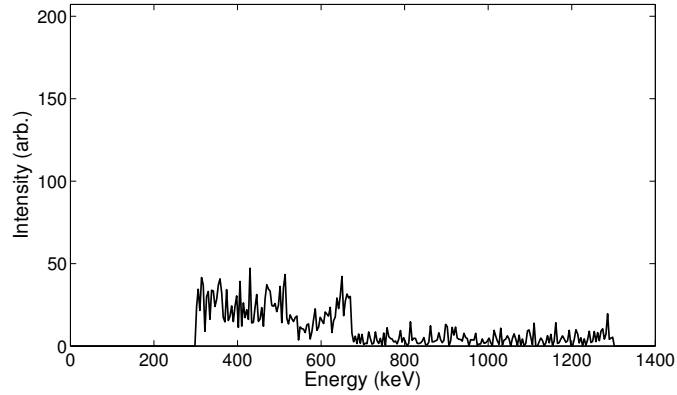
(a)
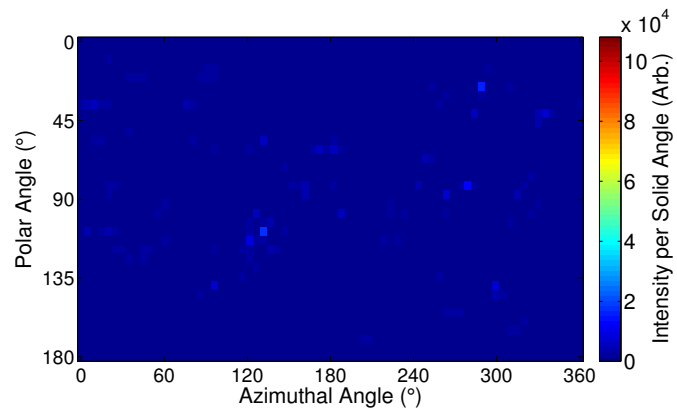
(b)
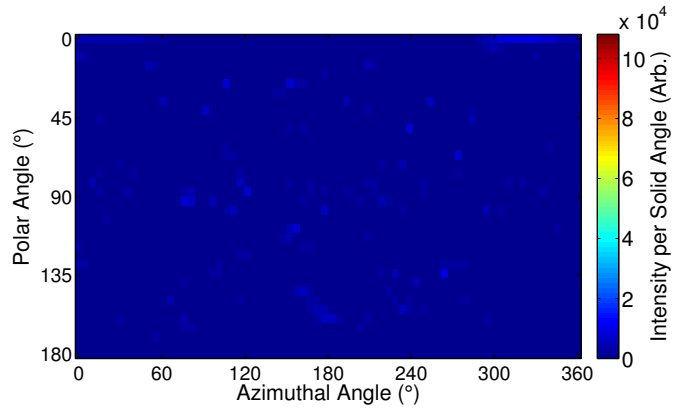
(c)

Figure 5.12: Results for the backdrop space after 20 iterations of the proposed motion-compensated EM algorithm for a ^{137}Cs source moving counterclockwise and a ^{22}Na source moving clockwise: (a) reconstructed incident energy spectrum for the entire backdrop space; (b) reconstructed spatial distribution for the ^{137}Cs energy slice; (c) reconstructed spatial distribution for the ^{22}Na energy slice.
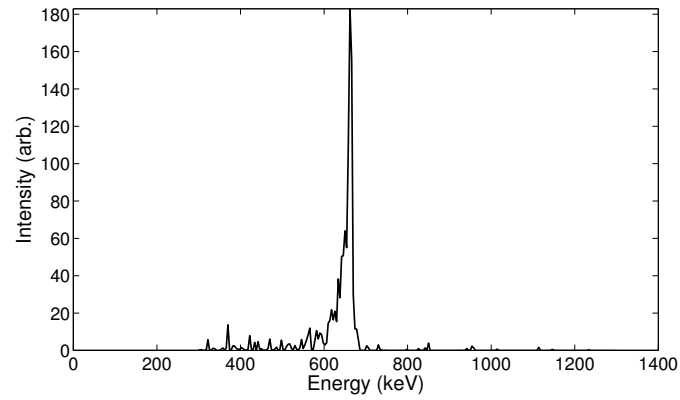
123

(a)

Intensity per Solid Angle (Arb.)

(b)   (c)   (d)   (e)
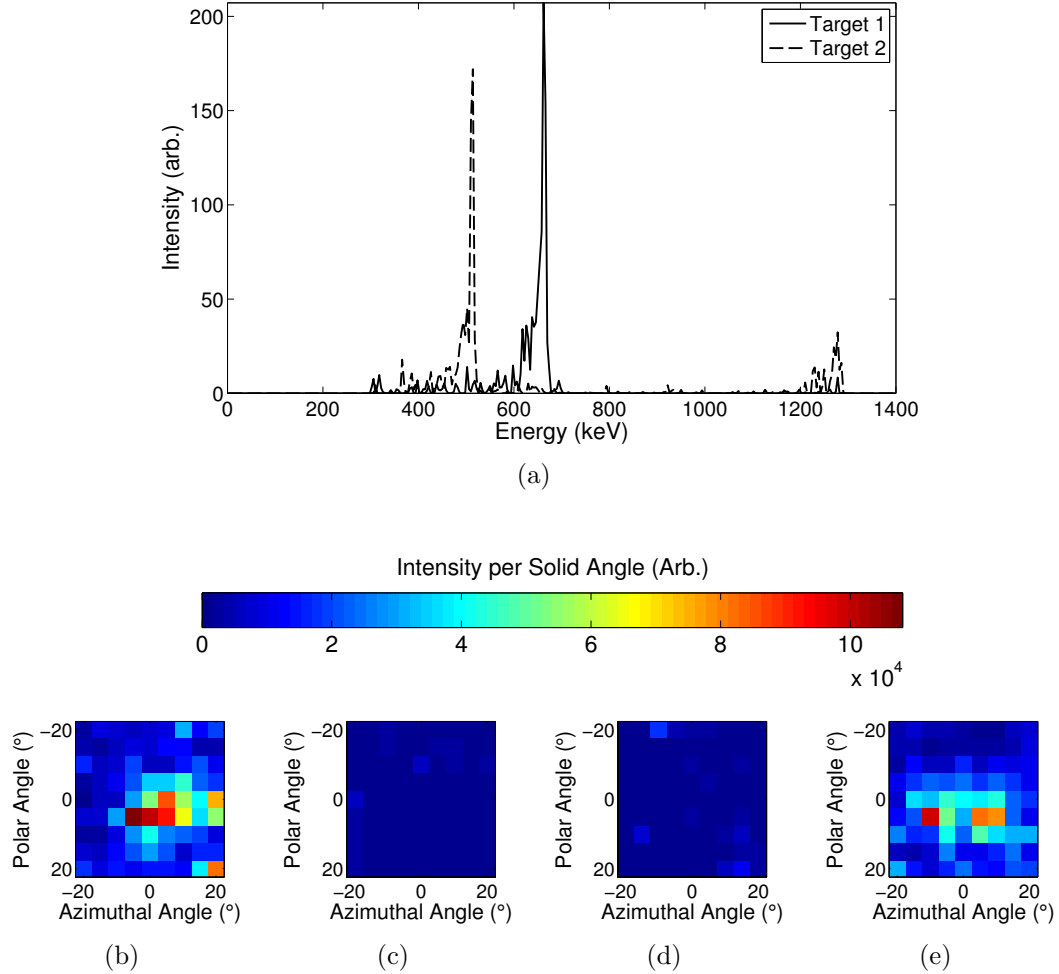
Figure 5.13: Results for the two target-object spaces after 20 iterations of the proposed motion-compensated EM algorithm for a ^{137}Cs source moving counterclockwise and a ^{22}Na source moving clockwise: (a) reconstructed incident energy spectrum for both targets; (b) reconstructed spatial distribution for the ^{137}Cs energy slice in Target 1; (c) reconstructed spatial distribution for the ^{22}Na energy slice in Target 1; (d) reconstructed spatial distribution for the ^{137}Cs energy slice in Target 2; (e) reconstructed spatial distribution for the ^{22}Na energy slice in Target 2.

the motion of the targets of interest are well-known. However, several improvements are key to make the algorithm more suitable for use in real-world scenarios.

In this work, each target had the same extent and the source was located at the center of the target. This assumption may not be realistic in a fielded device. For example, if the reconstruction is used at a border crossing, cars in the far lane will subtend a smaller solid angle than those in the near lane. Thus, the target domains used for the cars in the near lane should have a greater extent. When a larger extent is used for the target domain, the orientation of the objects in that domain also become more important. If the source is located at the center of the target, the orientation would not matter, but in most cases the source will not be in the center. Therefore, more careful treatment of the orientation of the target space may be needed for some applications.

Another assumption used in the current reconstruction model is that no occlusions are present between targets and the detector. In the border crossing scenario just described, this may be a terrible assumption if a loaded dump truck rolls in between the detector and a target of interest. However, if optical video cameras are used to track the targets, it may be relatively straightforward to use that information to estimate the attenuation between the detector and any given target as a function of time. This would complicate the model, but could drastically improve results in certain cases.

One last key point to consider is that in many cases, including the border crossing scenario, it is likely that more than one detector will be in use. This 2-D reconstruction has no way to benefit from multiple detectors at different locations. For that type of problem, this motion compensation algorithm should be applied to the 3-D domain. This adaptation of the reconstruction should be straightforward, but it could emphasize the importance of handling the orientation and attenuation more carefully.

CHAPTER VI

Image Reconstruction with Unknown Source Motion

Chapter V described a new algorithm that adds a set of spatial bins to the imaging domain in order to track moving objects. This alternative form of spatial binning allows one to simultaneously reconstruct a mixture of stationary and moving sources without blurring and with minimal cross contamination between sources. Another key aspect of the target spatial binning concept is that it essentially removes the time-dependence from the final solution. Although the direction associated with the target-image domain changes as a function of time to match the direction of the target, the result is simply composed of a backdrop image and a target image. There is not a set of target images as a function of time. This ability to reduce the dimensionality of the problem is desirable because it does not distribute the already scarce number of counts over many time bins.

However, sometimes the target spatial binning is impractical. In situations where the number of possible targets is large, like a crowd in an airport, it may be impractical to track all of them, especially if the targets are moving in and out of the optical field of view or if they are crossing paths. Even if it were possible to track them all, including them all in the image reconstruction may prove to be a challenge since many targets could have very similar paths. For these cases where there are too

many targets, the targets are difficult to track, or there are no known target paths, we must take a different approach.

This chapter discusses the explicit addition of the time domain in the image reconstruction. That is to say, here the reconstruction algorithms will estimate the radiation spatial distribution as a function of time and possibly energy. (Think radiation *movie*.) Three different methods are discussed that have different ways to handle the sparsity of the data and smoothing the image in the time domain.

6.1 Time-Dependent MLEM

The simplest and most straightforward way to include the time domain in the image reconstruction is to bin the incoming events by acquisition time and treat each time bin as a separate reconstruction. In this way, one gets a snapshot of the radiation distribution for each time interval in the binning structure. This approach is the most analogous to making a "radiation movie".

The addition of the time domain may sound similar to the addition of the energy domain as discussed when the complete system model was introduced in Section 3.3.2.4. However, there is a big difference between the time domain and the energy domain. For the energy domain, even if the system has perfect energy resolution (*i.e.,* the energy depositions are known exactly), it is still possible that the total energy deposited in a given event may not be the true energy of the incident photon. Thus, partial energy deposition is included in the system model which couples the space and energy domains, and the MLEM reconstruction estimates the most likely source distribution in space *and* energy.

However, assuming that all interactions resulting from a single incident photon occur simultaneously and ignoring deadtime and pulse pile-up effects, it is not possible for an energy deposition in the detector at time t_i to be caused by a photon incident on the system at some previous time t_{i-n} or some future time t_{i+n}. Therefore, the system

response functions described in Section 3.3.2 are still valid for this time-dependent reconstruction since we assume each time bin is completely independent of the others, and we use the timestamp of each event only to determine the time bin to which it belongs.

6.1.1 Implementation

In order to include the time domain in the MLEM reconstruction, we modify the parameterization of the source distribution $\lambda(\vec{r}, E, t)$ as follows:

$$\lambda(\vec{r}, E, t) = \sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} b_{lj}(\vec{r}, E, t) \tag{6.1}$$

where λ_{lj} represents the unknown intensity corresponding to the basis function b_{lj} which is now described by a simple rect function covering the jth spatial-energy bin and the lth time bin. As before, N_0 is the total number of spatial-energy bins for a given time and L is the total number of time bins.

Since the time-dependent list-mode log likelihood previously derived in Section 5.3.2 is still valid, we can rewrite it using the new parameterization of $\lambda(\vec{r}, E, t)$ by substituting (6.1) into (5.24):

$$\mathsf{L}(\lambda) =$$

$$\sum_{m=1}^{n} \log \left(\iint_{\Omega} \mathsf{p}(\boldsymbol{v}_m \mid \vec{r}, E, t_m, D_m) s(\vec{r}, E, t_m) \sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} b_{lj}(\vec{r}, E, t_m) \, \mathrm{d}\vec{r} \mathrm{d}E \right)$$

$$- \int_0^{\tau} \iint_{\Omega} s(\vec{r}, E, t) \sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} b_{lj}(\vec{r}, E, t) \, \mathrm{d}\vec{r} \mathrm{d}E \, \mathrm{d}t. \tag{6.2}$$

Now, rearrange and make simplifying substitutions

$$\mathsf{L}(\lambda) = \sum_{m=1}^{n} \log \left[\sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} u_{ljm} \right] - \sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} v_{lj} \tag{6.3}$$

where,

$$u_{ljm} = \iint\limits_{\Omega} \mathsf{p}(\boldsymbol{v}_m \mid \vec{r}, E, t_m, D_m) s(\vec{r}, E, t_m) b_{lj}(\vec{r}, E, t_m) \, \mathrm{d}\vec{r} \mathrm{d}E \qquad (6.4)$$

$$v_{lj} = \int\limits_{0}^{\tau} \iint\limits_{\Omega} s(\vec{r}, E, t) b_{lj}(\vec{r}, E, t) \, \mathrm{d}\vec{r} \mathrm{d}E \, \mathrm{d}t. \qquad (6.5)$$

Following the same derivation found in Section 5.3.3, the EM algorithm update equation for this simple time-binned EM reconstruction is

$$\lambda_{lj}^{i+1} = \frac{\lambda_{lj}^i}{v_{lj}} \sum_{m=1}^{n} \frac{u_{ljm}}{\sum_{l'=1}^{L} \sum_{j'=1}^{N_0} \lambda_{l'j'}^i u_{l'j'm}}. \qquad (6.6)$$

If the basis functions $b_{lj}(\vec{r}, E, t)$ were something other than simple non-overlapping rectangular bin functions, (6.6) is the final simplified update equation. However, since we are using these simple rectangle basis functions, (6.6) can be simplified to:

$$\lambda_{lj}^{i+1} = \frac{\lambda_{lj}^i}{v_{lj}} \sum_{m=1}^{n} \frac{u_{ljm}}{\sum_{j'=1}^{N_0} \lambda_{lj'}^i u_{lj'm}} \qquad (6.7)$$

since the system response $u_{l'j'm}$ in (6.6) will have zero value for any time bin other than l.

The ability to simplify from (6.6) to (6.7) makes the computation faster and easier to parallelize. The computation is faster because the outer loop over the time bins has been removed from the denominator of the update equation. More importantly this simplification allows one to calculate the solution for each time bin l independently using different processing threads. It also saves memory (if the system matrix is stored in memory) since each time bin can have its own system matrix which has a dimension of n_l rows by N_0 columns. In contrast, one large system matrix would have n rows by $N_0 \cdot L$ columns, where n_l is the number of recorded counts in time bin l. However, memory would be less of an issue if sparse matrix storage is used.

129

6.1.2   Performance

We used a simple measurement of a 30 μCi ^{137}Cs source moving around the detector array to demonstrate the capability of the time-dependent MLEM reconstruction. During the experiment, the source moved at a relatively slow and constant speed, and the distance from the detector was constant at about 1 m. We used 30 time bins to evenly discretize the time domain for the six minute measurement. For simplicity, we used the standard full-energy assumption MLEM model and only reconstructed photopeak events.

The reconstructed time distribution as a function of direction is shown in Figure 6.1. The intensity peaks at the time that the source crosses the direction corresponding to each distribution. In this case, there are enough counts (about 5700) and few enough time bins such that the distributions show little statistical noise from the reconstruction. By comparison, a similar plot made with results from the SBP reconstruction would show a broad peak with large tails because of the large point spread function of the SBP reconstruction.

Figure 6.2 shows the reconstructed spatial distribution as a function of time. As with the time domain, there is little statistical noise in the reconstructed images. The source location is easily identifiable with minimal artifacts. These results are expected since, for this simple time-binned reconstruction, each time bin is essentially a standard MLEM reconstruction without time-dependence.

The previous results show that the simple addition of the time domain works well when there are plenty of counts in each time bin to allow the reconstruction to generate an accurate estimate of the source direction with little statistical noise. However, in many cases the source strength may be weaker and/or the source may be moving faster. In this case where only a handful of counts may be recorded in a given time window, the standard MLEM reconstruction begins to reproduce noisy images from which it is difficult to discern the source direction.

130
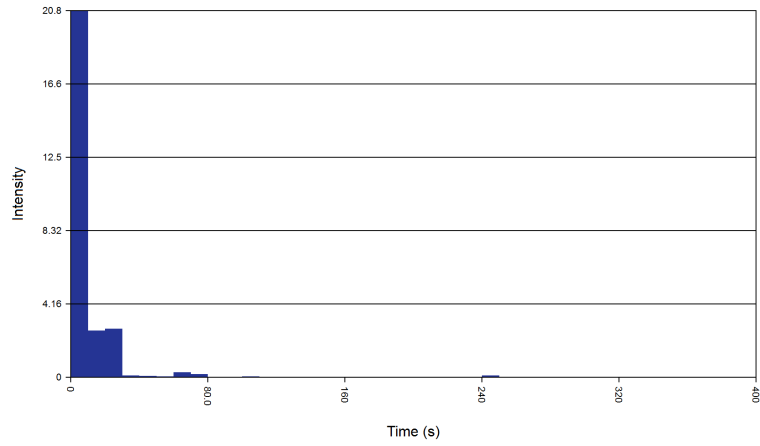
(a)
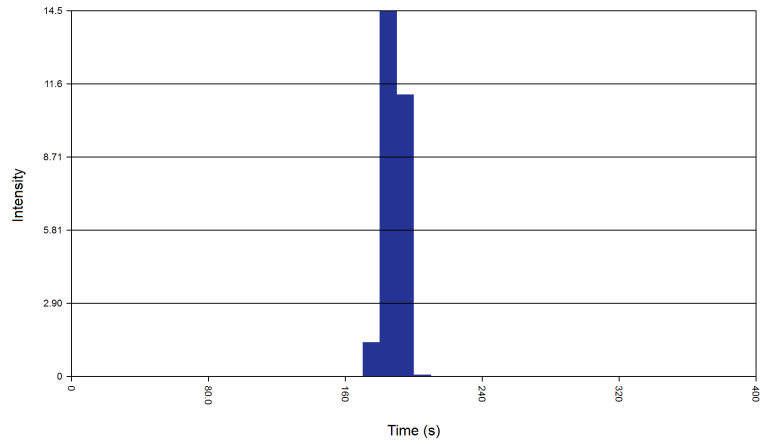
(b)
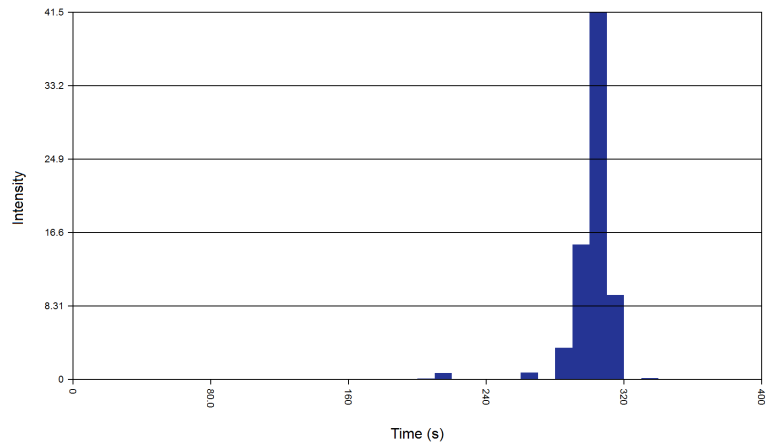
(c)

Figure 6.1: Unregularized MLEM solution after 25 EM iterations for the time domain in the direction of the hotspot shown in: (a) Figure 6.2(a), (b) Figure 6.2(b), and (c) Figure 6.2(c).

131
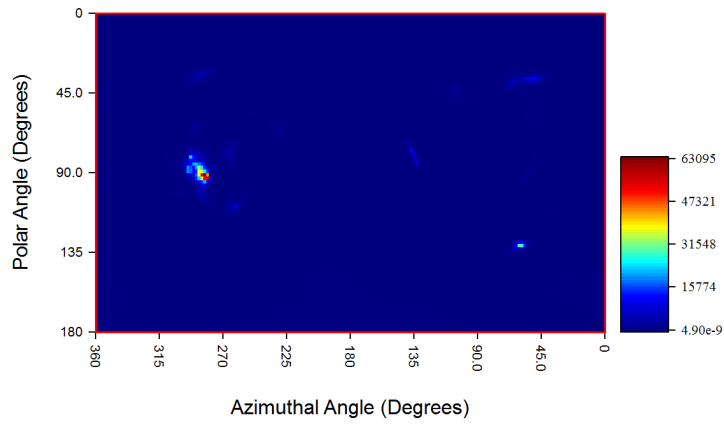
(a)
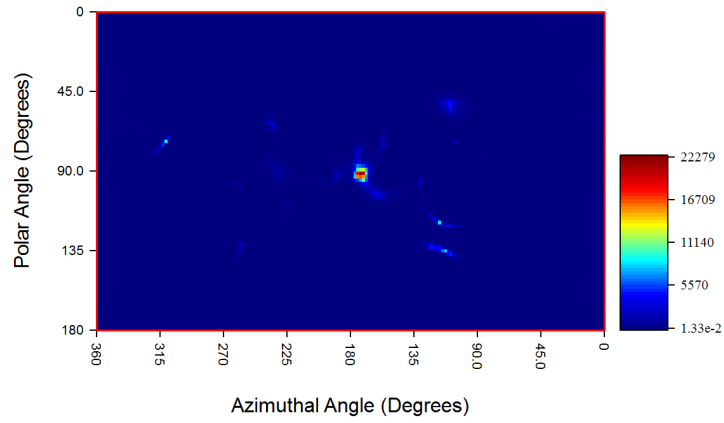
(b)
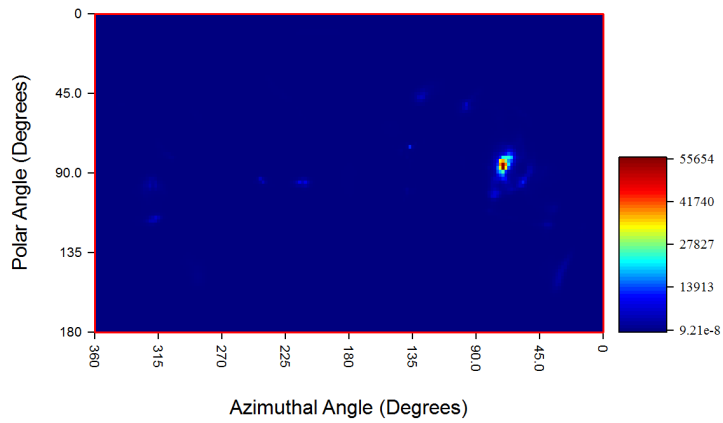
(c)

Figure 6.2: Unregularized MLEM solution after 25 EM iterations in the spatial domain for the time corresponding the peak in: (a) Figure 6.1(a), (b) Figure 6.1(b), and (c) Figure 6.1(c).

132

We performed another experiment using the same apparatus as the experiments for the known motion case in Section 5.4. In this experiment a single ^{137}Cs source rotated around the detector array over the course of about 40 mins. We downsampled the data so that there are about 3000 total imageable photopeak counts, and then we used the downsampled dataset both to compare the time-dependent algorithms presented here and for the optimization study at the end of the chapter.

To test the limits of the standard MLEM algorithm, we reconstruct this dataset using 300 time bins. This binning results in approximately 10 counts per time window. Figure 6.3 shows the reconstructed time distribution after 25 EM iterations as a function of direction. With the increased number of time bins, the time at which the source passed each direction is still discernible, but the statistical noise plays a much more dominant role. The same is true for the spatial distribution seen in Figure 6.4. Although it may be difficult to tell with just the few snapshots in time displayed here, the overall path of the source can be determined by eye from the reconstructed images with careful examination. However, many artifacts are now present in the distributions making it difficult to know the true direction at any given time, especially if one is just looking at individual snapshots like in Figure 6.4

To overcome the effects of low statistics for low activity or fast moving sources, we have have developed new reconstruction methods.

6.2 MLEM with Time Smoothing Regularizer

In a standard (unregularized) maximum-likelihood reconstruction, the goal is to maximize the agreement between the estimated source distribution and the recorded measurements. However, sometimes one may have previous knowledge about how the distribution should behave. In this case, instead of just using the likelihood function, we use an objective function, Ψ, of the following form to solve for the
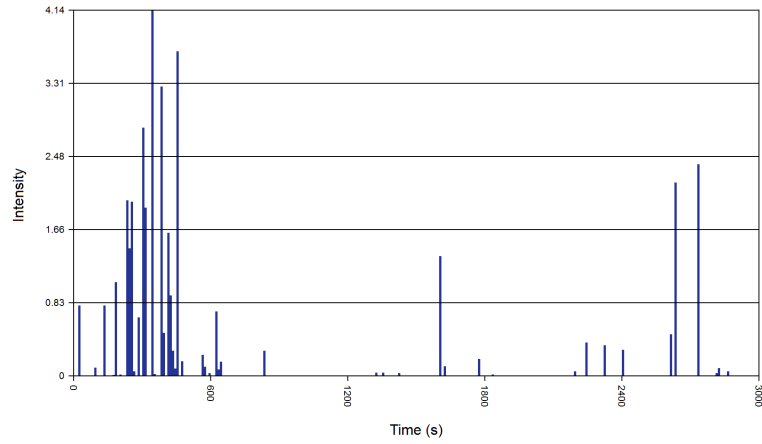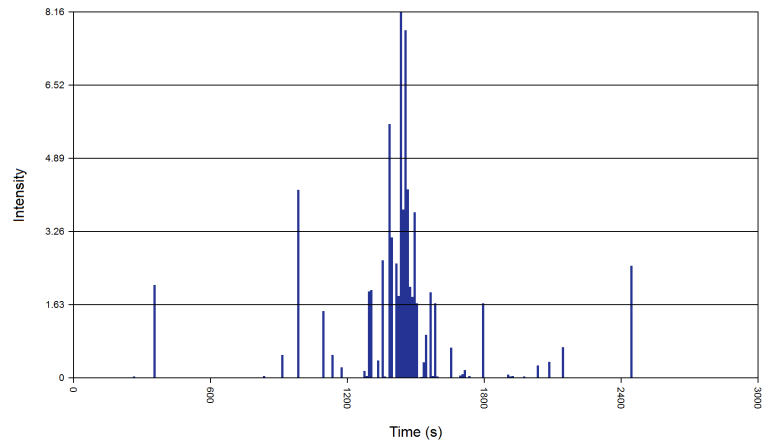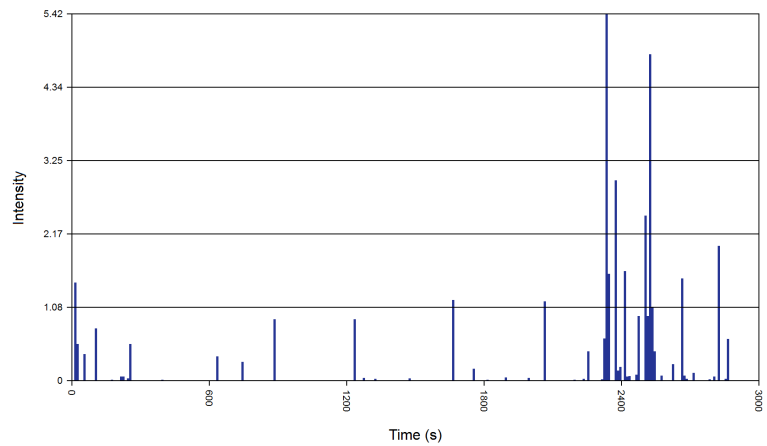
Figure 6.3: Unregularized MLEM solution after 25 EM iterations. The plots show the source intensity as a function of time along the equator at an azimuthal angle of (a) 45°, (b) 180°, and (c) 315°.
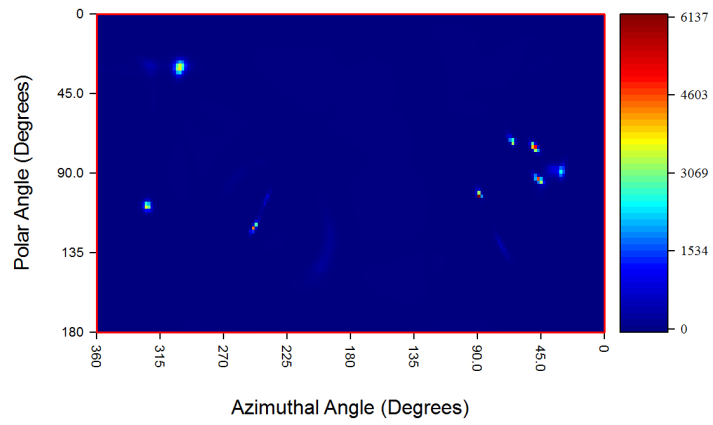
(a)
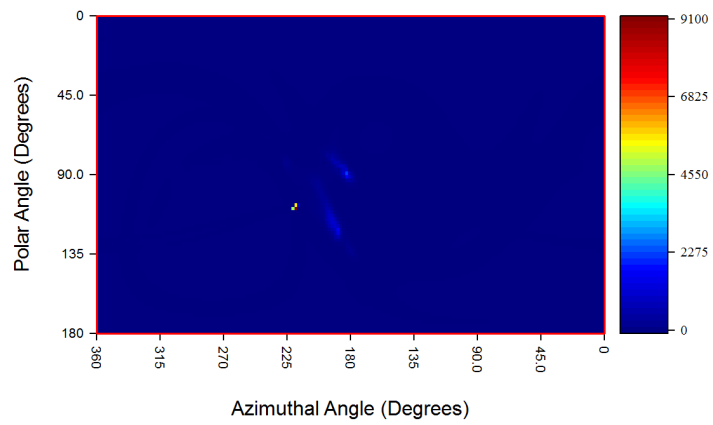
(b)
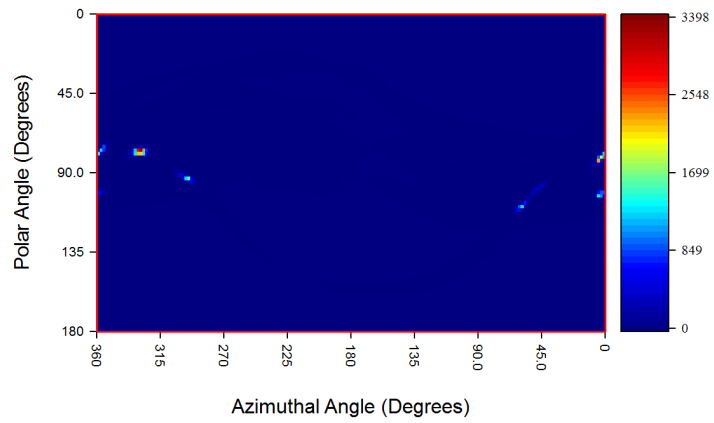
(c)

Figure 6.4: Unregularized MLEM solution after 25 EM iterations. The figures show the spatial source distribution at (a) 230 s, (b) 1500 s, and (c) 2800 s.

source distribution

$$\hat{\lambda} = \arg\max_{\lambda} \ \Psi(\lambda) \tag{6.8}$$

$$\Psi(\lambda) = \mathsf{L}(\lambda) - \delta\mathsf{R}(\lambda) \tag{6.9}$$

where $\mathsf{R}(\lambda)$ is the penalty function that lowers the objective function if λ does not have the expected form, and the regularization parameter δ is used to determine how strongly to enforce the regularization. $\mathsf{R}(\lambda)$ may take many different forms depending on the desired effect. For example, Lingenfelter [63] describes several different penalty functions for encouraging spatial sparsity.

However, we want to encourage the solution to be smooth in the time domain since we assume that the sources are moving relatively slowly compared to the size of the time bins. For this case we use a simple roughness penalty function [64]

$$\mathsf{R}(\lambda) = \frac{1}{2}\sum_{l=1}^{L}\sum_{j=1}^{N_0}(\lambda_{lj} - \lambda_{l-1\,j})^2. \tag{6.10}$$

This function will discourage the intensity of image pixel j from changing greatly between time bins l and $l-1$. Note that the sum over l goes from 1 to L so that we can combine terms later, and the boundary conditions are defined after the simplification such that all of the subscripts are valid.

6.2.1 Derivation

Using the parameterized time-dependent log likelihood (6.3) and the penalty function (6.10), the objective function is

$$\Psi(\lambda) = \sum_{m=1}^{n} \log \left[\sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} u_{ljm} \right] - \sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} v_{lj}$$
$$-\frac{\delta}{2} \sum_{l=1}^{L} \sum_{j=1}^{N_0} (\lambda_{lj} - \lambda_{l-1\,j})^2. \tag{6.11}$$

Now we want to find the EM update equation to solve for the $\hat{\lambda}$ that will maximize this objective function. To do that we will follow the derivation found in Section 5.3.3.

First, expand the penalty term and combine like sums

$$\Psi(\lambda) = \sum_{m=1}^{n} \log \left[\sum_{l=1}^{L} \sum_{j=1}^{N_0} \lambda_{lj} u_{ljm} \right]$$
$$-\sum_{l=1}^{L} \sum_{j=1}^{N_0} \left[\lambda_{lj} v_{lj} + \frac{\delta}{2} \left(\lambda_{lj}^2 + \lambda_{l-1\,j}^2 - 2\lambda_{lj}\lambda_{l-1\,j} \right) \right] \tag{6.12}$$

Next introduce β_{ljm} and rewrite (6.12)

$$\Psi(\lambda) = \sum_{m=1}^{n} \log \left[\sum_{l=1}^{L} \sum_{j=1}^{N_0} \frac{\beta_{ljm} \lambda_{lj} u_{ljm}}{\beta_{ljm}} \right]$$
$$-\sum_{l=1}^{L} \sum_{j=1}^{N_0} \left[\lambda_{lj} v_{lj} + \frac{\delta}{2} \left(\lambda_{lj}^2 + \lambda_{l-1\,j}^2 - 2\lambda_{lj}\lambda_{l-1\,j} \right) \right] \tag{6.13}$$

where

$$\beta_{ljm} > 0 \quad \forall \ l, j, m \qquad \text{and} \qquad \sum_{l=1}^{L} \sum_{j=1}^{N_0} \beta_{ljm} = 1 \quad \forall \ m$$

Remember that $\log(x)$ is a concave function (5.30) under the conditions (5.31), and therefore a new surrogate function can be written that is separable in λ_{lj} and is no

greater than the original objective function

$$\Psi(\lambda) \geq \Phi(\lambda; \beta_{ljm}) = \sum_{l=1}^{L} \sum_{j=1}^{N_0} \phi_{lj}(\lambda_{lj}; \beta_{ljm}) \qquad (6.14)$$

where

$$\phi_{lj}(\lambda_{lj}; \beta_{ljm}) = \sum_{m=1}^{n} \beta_{ljm} \log \left[\frac{\lambda_{lj} u_{ljm}}{\beta_{ljm}} \right]$$
$$- \lambda_{lj} v_{lj} - \frac{\delta}{2} \left(\lambda_{lj}^2 + \lambda_{l-1\,j}^2 - 2\lambda_{lj}\lambda_{l-1\,j} \right). \qquad (6.15)$$

Updating the β_{ljm} term is performed exactly the same as it was before, namely

$$\beta_{ljm}^i = \frac{\lambda_{lj}^i u_{ljm}}{\sum_{l'=1}^{L} \sum_{j'=1}^{N_0} \lambda_{l'j'}^i u_{l'j'm}} = \frac{\lambda_{lj}^i u_{ljm}}{\sum_{j'=1}^{N_0} \lambda_{lj'}^i u_{lj'm}}. \qquad (6.16)$$

Plug the current β_{ljm}^i term into (6.15) to set the current surrogate function. Then, previously we maximized $\phi_{lj}(\lambda_{lj}; \beta_{ljm}^i)$ to find λ_{lj}^{i+1}; however, now because the penalty term is coupled across the time domain we must maximize $\Phi(\lambda^i; \beta_{ljm})$ with respect to λ_{lj} instead

$$\frac{\partial}{\partial \lambda_{lj}} \Phi(\lambda; \beta_{ljm}^i) = 0 = \sum_{m=1}^{n} \frac{\beta_{ljm}^i}{\lambda_{lj}} - v_{lj} - \delta\left(2\lambda_{lj} - \lambda_{l-1\,j} - \lambda_{l+1\,j}\right). \qquad (6.17)$$

Now multiply both sides by λ_{lj}

$$0 = \left(-2\delta\right)\lambda_{lj}^2 + \left(\delta\lambda_{l-1\,j} + \delta\lambda_{l+1\,j} - v_{lj}\right)\lambda_{lj} + \sum_{m=1}^{n} \beta_{ljm}^i \qquad (6.18)$$

and solve for the new estimate of λ_{lj}

$$\lambda_{lj}^{i+1} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \Rightarrow \frac{-b - \sqrt{b^2 - 4ac}}{2a} \qquad (6.19)$$

where

$$a = \begin{cases} -2\delta & 1 < l < L \\ -\delta & \text{otherwise} \end{cases}$$

$$b = \begin{cases} \delta\lambda_{l-1\,j}^{i+1} + \delta\lambda_{l+1\,j}^{i+1} - v_{lj} & 1 < l < L \\ \delta\lambda_{l+1\,j}^{i+1} - v_{lj} & l = 1 \\ \delta\lambda_{l-1\,j}^{i+1} - v_{lj} & l = L \end{cases}$$

$$c = \sum_{m=1}^{n} \beta_{ljm}^i = \sum_{m=1}^{n} \frac{\lambda_{lj}^i u_{ljm}}{\sum_{j'=1}^{N_0} \lambda_{lj'}^i u_{lj'm}}.$$

Notice that there will always be a positive and negative real root since δ and β_{ljm} are greater than zero, but only the positive root is taken since λ can only take on positive values. Also, at the boundaries, a one-sided difference is taken for the penalty term.

6.2.2   Implementation

The update equation derived in the previous section takes on a deceivingly simple form. At first glance it appears that the update can be found by simply plugging in the terms a, b, and c into (6.19). However, after closer inspection, the b term in (6.19) depends on values of λ in the current iteration. This dependency complicates the implementation slightly because the solution is coupled in the time domain and each spatial-energy bin j must be solved simultaneously over all the time bins.

Green [65] suggests using the "one-step-late" algorithm to avoid this implementation problem. This algorithm uses the solution of the previous iteration, λ_{lj}^i, for the penalty term in the objective function. Note that this is a somewhat heuristic approach and does not exactly compute the true penalized-likelihood solution, but it simplifies the implementation. Start by setting the derivative of the objective function

to zero (6.17)

$$0 = \sum_{m=1}^{n} \frac{\beta_{ljm}^{i}}{\lambda_{lj}} - v_{lj} - \delta \big(2\lambda_{lj}^{i} - \lambda_{l-1\,j}^{i} - \lambda_{l+1\,j}^{i} \big). \tag{6.20}$$

Now rearrange terms and substitute β_{ljm}^{i} using (6.16)

$$v_{lj} + \delta \big(2\lambda_{lj}^{i} - \lambda_{l-1\,j}^{i} - \lambda_{l+1\,j}^{i} \big) = \frac{\lambda_{lj}^{i}}{\lambda_{lj}} \sum_{m=1}^{n} \frac{u_{ljm}}{\sum_{j'=1}^{N_0} \lambda_{lj'}^{i} u_{lj'm}} \tag{6.21}$$

and solve for the new estimate of λ_{lj}

$$\lambda_{lj}^{i+1} = \frac{\lambda_{lj}^{i}}{v_{lj} + \delta \big(2\lambda_{lj}^{i} - \lambda_{l-1\,j}^{i} - \lambda_{l+1\,j}^{i} \big)} \sum_{m=1}^{n} \frac{u_{ljm}}{\sum_{j'=1}^{N_0} \lambda_{lj'}^{i} u_{lj'm}}. \tag{6.22}$$

Green proved that this algorithm does converge and that it will converge relatively quickly if the smoothing parameter δ is small. However, in our case sometimes the smoothing required will be large. Also, with this implementation there is a chance that the penalty term will cause the solution to be undefined or negative. Thus, there is a limited range of δ values that can be chosen, and one would have to be mindful of that range when using the algorithm.

Because of these shortcomings, we propose a different implementation. Notice that the β_{ljm}^{i} values are independent of each other in the time domain. Thus, we solve (6.19) iteratively (always using the most current estimate of λ_{lj}) without having to recalculate β_{ljm}^{i} after each (sub-)iteration

$$\lambda_{lj}^{i+s/S} = \frac{-b - \sqrt{b^2 - 4ac}}{2a} \tag{6.23}$$

Algorithm VI.1 Pesudocode for penalized-likelihood EM update.

for $i = 1 \rightarrow numIterations$ **do**
  Set current surrogate function
    for all l **do**
      for all j **do**
        Calc β^i_{ljm} (6.16)
      end for
    end for
  Iteratively solve for new estimate of source distribution
    for $s = 1 \rightarrow numSubiterations$ **do**
      for all j **do**
        for all l **do**
          Calc $\lambda^{i+s/numSubIterations}_{lj}$ (6.19)
        end for
      end for
    end for
end for

where s is the current sub-iteration, S is the total number of sub-iterations and

$$
a = \begin{cases} -2\delta & 1 < l < L \\ \\ -\delta & \text{otherwise} \end{cases}
$$

$$
b = \begin{cases} \delta\lambda^{i+s/S}_{l-1\,j} + \delta\lambda^{i+(s-1)/S}_{l+1\,j} - v_{lj} & 1 < l < L \\ \delta\lambda^{i+(s-1)/S}_{l+1\,j} - v_{lj} & l = 1 \\ \delta\lambda^{i+s/S}_{l-1\,j} - v_{lj} & l = L \end{cases}
$$

$$
c = \sum_{m=1}^{n} \beta^i_{ljm} = \sum_{m=1}^{n} \frac{\lambda^i_{lj} u_{ljm}}{\sum_{j'=1}^{N_0} \lambda^i_{lj'} u_{lj'm}}
$$

In essence, after each main iteration where the surrogate function is updated, the penalty term is used to smooth out the time domain iteratively. Algorithm VI.1 shows some pseudocode that describes the algorithm. The key here is that the sub-iterations do not significantly slow down the computation time because it is a simple root calculation without the large sums found in the main iterations.

    Another way to speed up the algorithm would be to calculate the sub-iterations

in parallel. Some care has to be taken to ensure that adjacent time bins are not being updated at the same time, since the solution at each time depends on the value of the adjacent time bins. A simple implementation is to update the odd time bins in parallel first then update the even time bins in parallel and repeat. This algorithm is similar to other block or grouped updates previously developed [66, 67], and because each step in the algorithm increases the penalized likelihood by design, one can show that the algorithm will converge [68].

6.2.3 Performance

We use the same low-count-rate dataset used to test the performance of the standard time-dependent MLEM reconstruction in Section 6.1.2 to demonstrate the capability of this penalized reconstruction. Figure 6.5 shows the reconstructed time distribution of the source intensity as a function of direction after 25 iterations. For comparison we use the same binning structure for this reconstruction, and it is immediately obvious that the regularization term has successfully smoothed the distribution. Some noise is still present, but the high-intensity noise spikes have been eliminated.

Figure 6.6 shows the reconstructed spatial distribution as a function of time. As with the time distribution, artifacts from noise are still present in the images, but the distributions are smoother, and as a result the hotspot tracks the source direction more consistently and is more easily identified by the eye.

The penalization factor, δ, greatly affects the resulting image quality and must be chosen by the user before the reconstruction can be performed. The results in the previous section were generated using a penalization factor of 0.1 which we selected after trying several other values. Figure 6.7 shows the reconstructed spatial distribution for a variety of penalization factors. If the penalty factor is too low like in Figure 6.7(a), the reconstructed distribution begins to resemble the standard time-
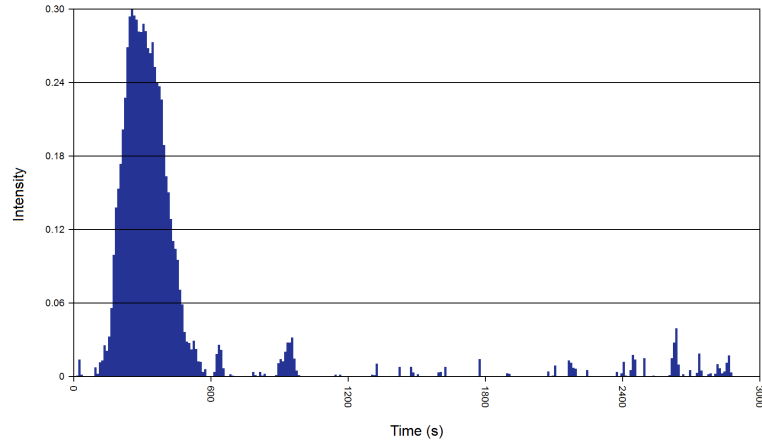
(a)
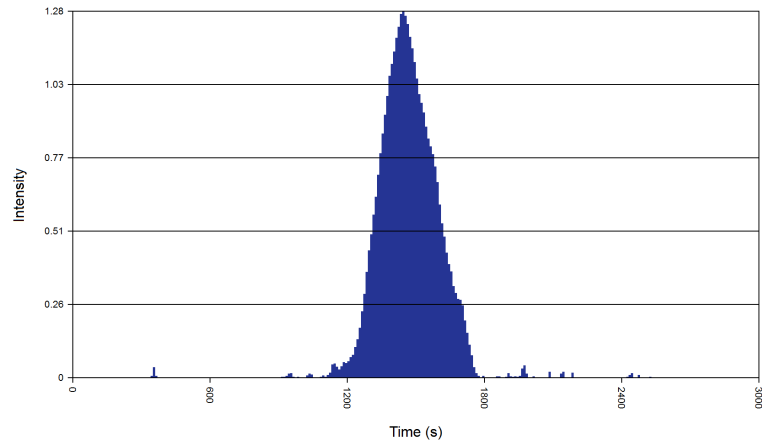
(b)
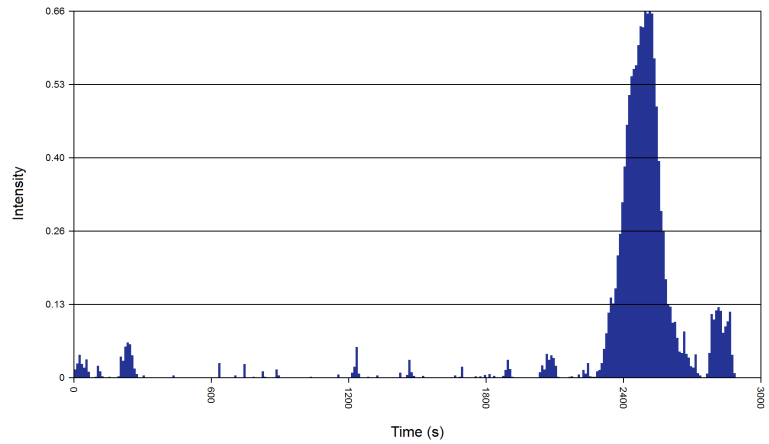
(c)

Figure 6.5: Regularized MLEM solution after 25 EM iterations using a penalization parameter of 0.1. The plots show the source intensity as a function of time along the equator at an azimuthal angle of (a) 45°, (b) 180°, and (c) 315°.
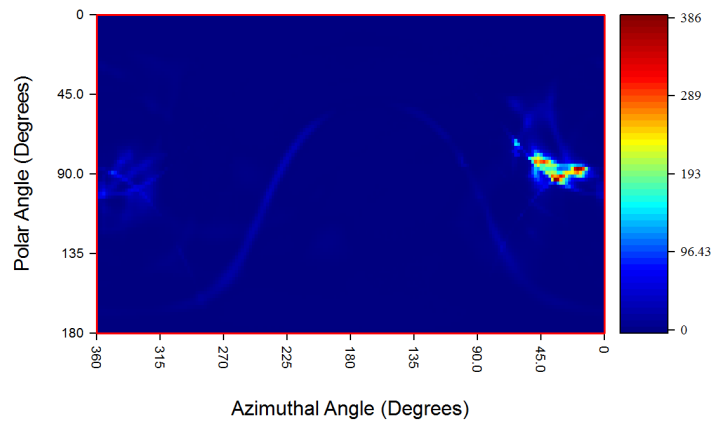
(a)
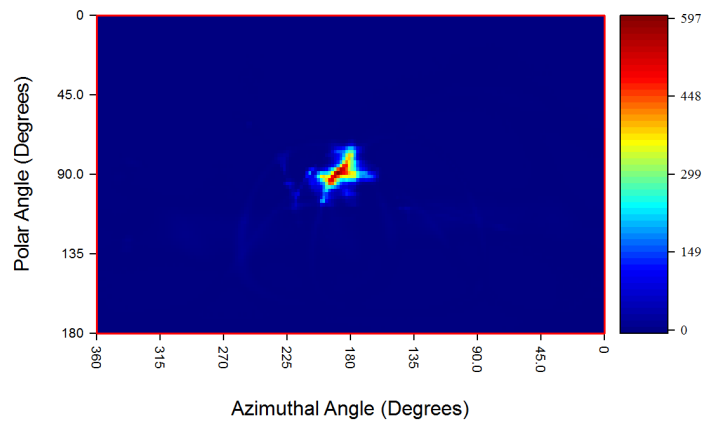
(b)
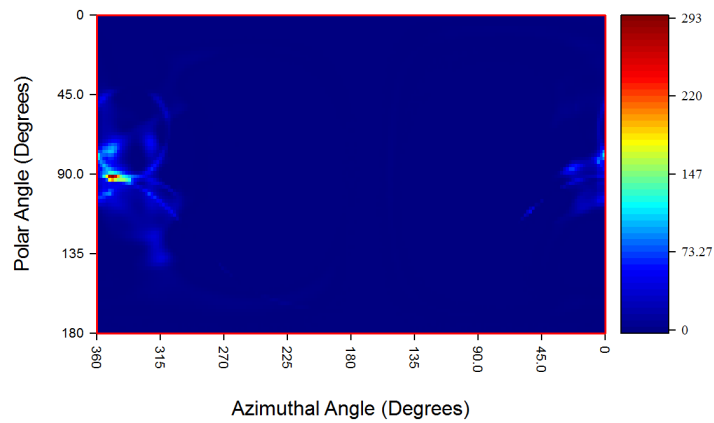
(c)

Figure 6.6: Regularized MLEM solution after 25 EM iterations using a penalization factor of 0.1. The figures show the spatial source distribution at (a) 230 s, (b) 1500 s, and (c) 2800 s.

dependent MLEM solution where noise artifacts dominate the solution. However, if the penalty factor is too large like in Figures 6.7(b) and 6.7(c), the reconstructed distribution becomes blurry. Thus, this parameter must be selected with care to ensure the best possible reconstruction.

6.3   Filtered Time-Dependent MLEM

This section describes an alternative way to address the poor results attained using the standard time-dependent MLEM algorithm described in Section 6.1. In the previous section, we added a penalty term to the objective function to smooth out the solution in the time domain. Here we use the solution from the standard time-dependent MLEM algorithm with standard 1-D digital filtering in the time domain to attempt to smooth or remove the statistical noise.

6.3.1   Average Filter

The first filter we consider is a simple linear filter that essentially calculates the moving average of the input signal. The discrete form of this filter is

$$
\lambda_{lj} = \begin{cases}
\displaystyle\sum_{w=-X}^{X} h\lambda_{l-wj} & l - X > 0 \quad \& \quad l + X \leq L \\[2ex]
\displaystyle\sum_{w=0}^{X} h\lambda_{l-wj} & l - X \leq 0 \\[2ex]
\displaystyle\sum_{w=-X}^{0} h\lambda_{l-wj} & l + X > L
\end{cases}
$$

$$
h = \begin{cases}
\frac{1}{W} & l - X > 0 \quad \& \quad l + X \leq L \\[2ex]
\frac{1}{l+X} & l - X \leq 0 \\[2ex]
\frac{1}{(L-l)+X} & l + X > L
\end{cases} \tag{6.24}
$$
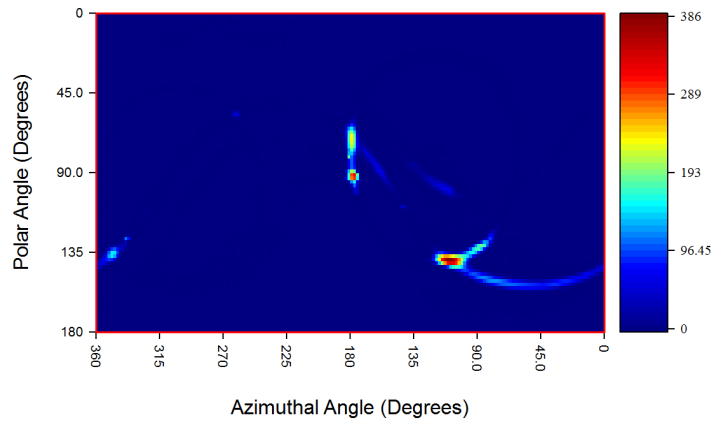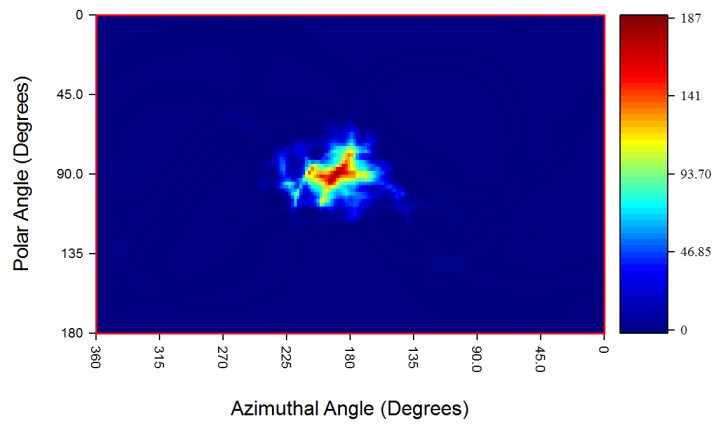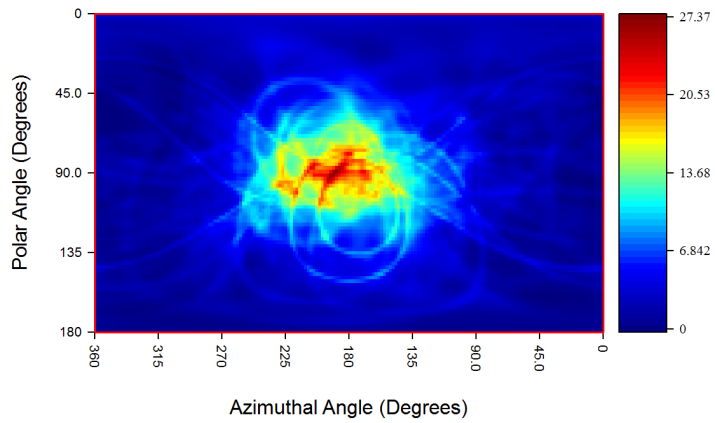
Figure 6.7: Source distribution at t=1500s for penalization factors of (a) δ=0.001, (b) δ=1.0, and (c) δ=100.0

where

$$X = \frac{W - 1}{2}.$$ (6.25)

W is the width of the filter and is a parameter that must be preselected similar to the δ parameter in the penalized likelihood method. W is usually chosen to be an odd value so that the smoothing is symmetric about each time bin, but an even W can be accommodated by rounding up both limits of the sum in (6.24). Notice that at the boundaries the averaging window shrinks to only include values that are part of the distribution. No zero padding or other artificial data is added to the ends of the distribution when averaging.

Figure 6.8 shows an example of what the filter does for an impulse input with different values of W. For impulse noise, the filter flattens and stretches the signal in time, and for a square signal, it flattens and stretches the signal and blurs the edges. The reduction in the amplitude of the noise is desirable, but the filter does not entirely remove the noise, and it distorts the edges of the signal which could potentially cause temporal blurring in the reconstructed image.

Figures 6.9 and 6.10 show the same unregularized results from Figures 6.3 and 6.4 after they are passed through the average filter with a window 21 bins wide. The filtered time domain results show a reduction of the high frequency noise seen in the unfiltered results. However, artifacts are still present which makes it difficult to discern the time at which the source passed a given direction. The spatial distribution results show similar traits. The overall source direction is easier to determine by eye, but even more artifacts are present in these distributions compared to the originals because the average filter has smoothed the high frequency noise through time. So even though the filter has reduced the intensity of the artifacts, it causes the influence of any given single spike of noise to be spread out over many time bins.
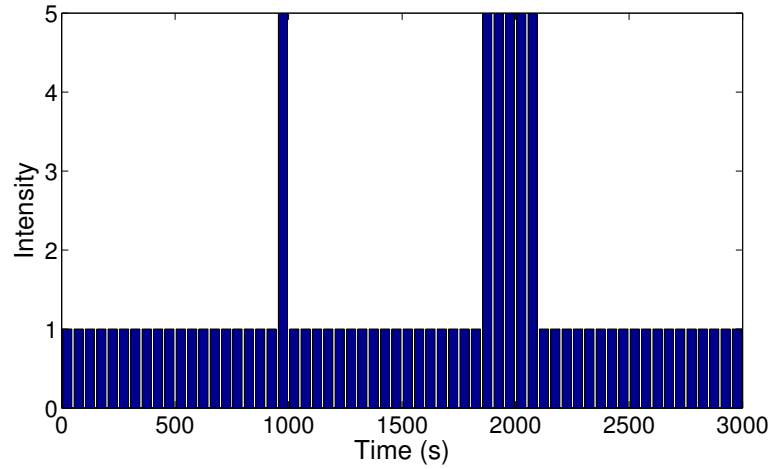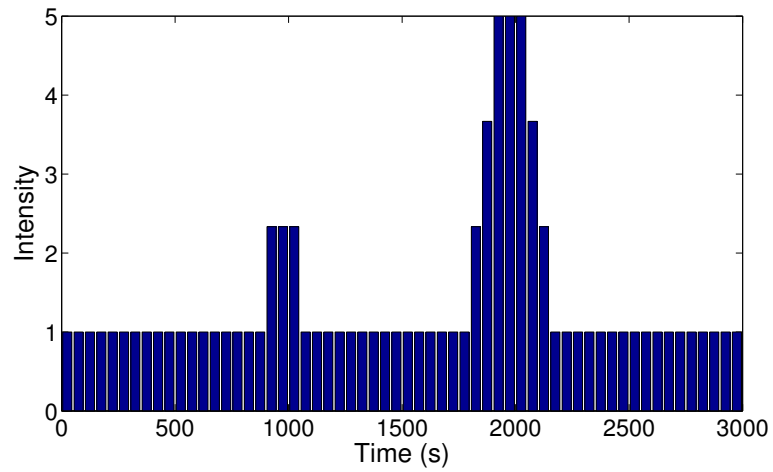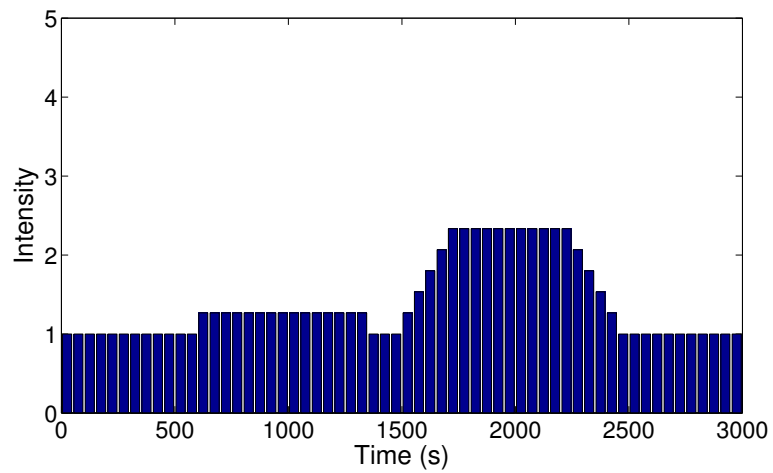
Figure 6.8: Example output from the average filter using the input signal shown in (a) using a window 3 bins wide (b) and 15 bins wide (c).

Figure 6.9: Unregularized MLEM solution after 25 EM iterations post-processed using the average filter with a window of 21 bins. The plots show the source intensity as a function of time along the equator at an azimuthal angle of (a) 45°, (b) 180°, and (c) 315°.

(a)

(b)

(c)

Figure 6.10: Unregularized MLEM solution after 25 EM iterations post-processed using the average filter with a window of 21 bins. The figures show the spatial source distribution at (a) 230 s, (b) 1500 s, and (c) 2800 s.

6.3.2 Median Filter

The average filter discussed in the previous section is a simple linear filter that performs a "smoothing" somewhat similar to the effect of the penalty term in the penalized-likelihood reconstruction. Since most of the artifacts seen in the standard time-dependent MLEM are sharp impulses uncorrelated in space or time, a simple average helps to lower the intensity of the artifacts, but it does not effectively remove them. A median filter is more appropriate for this "salt and pepper" type of noise. The median filter is non-linear, but it is still simple to implement and takes the following form

$$\lambda_{lj} = \text{median}[\lambda_{l-W\,j} \ldots \lambda_{lj} \ldots \lambda_{l+W\,j}] \tag{6.26}$$

where the filter window W can be odd or even. The filter is implemented by sorting all of the values in the filter window and then uses the value at the center of the sorted list as the filtered result. If W is even, the median value is calculated by taking the average of the two values splitting the median. Again, the filter does not use zero padding at the boundaries to fill the filter window. Figure 6.11 shows the result of using the median filter on the same distribution used to demonstrate the average filter. For this simple case, the impulse noise is completely removed for all choices of W, and the square signal is left unaltered except when the window becomes very large at $W = 15$. These traits are desirable, especially the complete removal of the impulse noise, since most of the time artifacts produced by low counting statistics are manifested as isolated peaks in the time domain.

Figures 6.12 and 6.13 show the same unregularized results from Figures 6.3 and 6.4 after they are passed through the median filter with a window 21 bins wide. It is obvious from Figure 6.12 that the median filter is a better fit for this type of noisy data. The filter has eliminated the artifacts far from the time at which the source passed each direction, and it did not drastically smooth the peaks. The filter has also
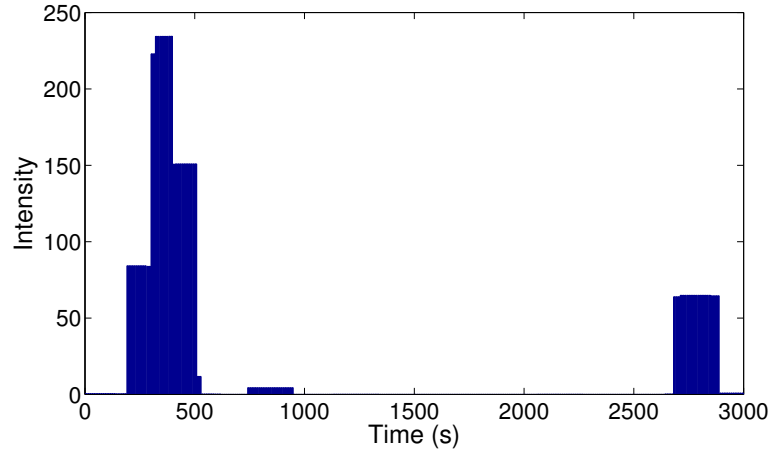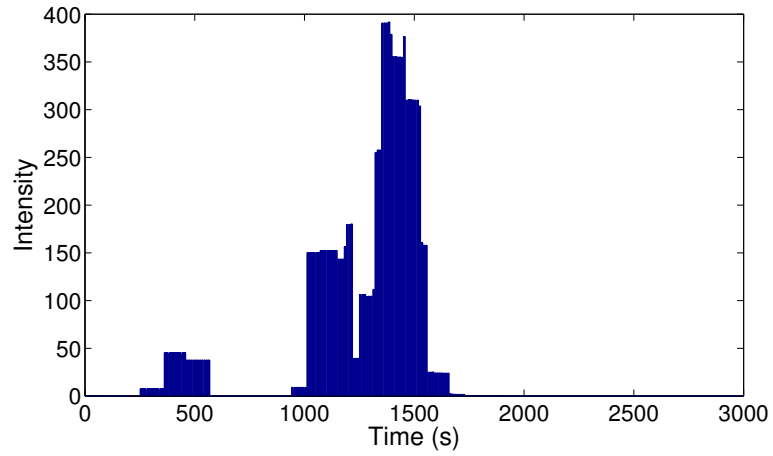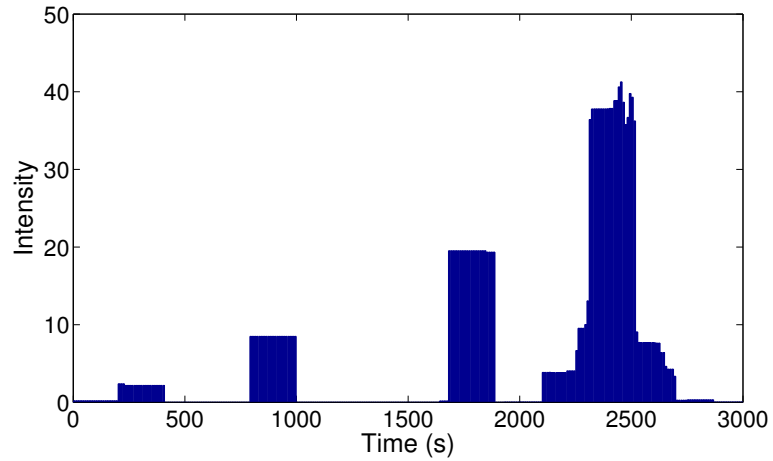
151
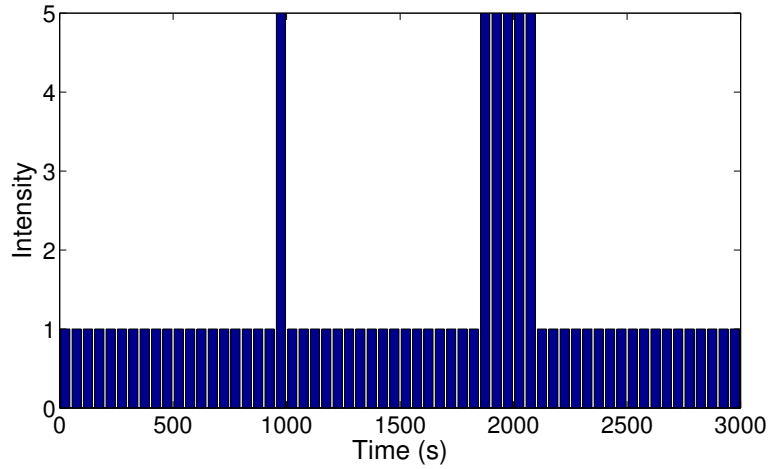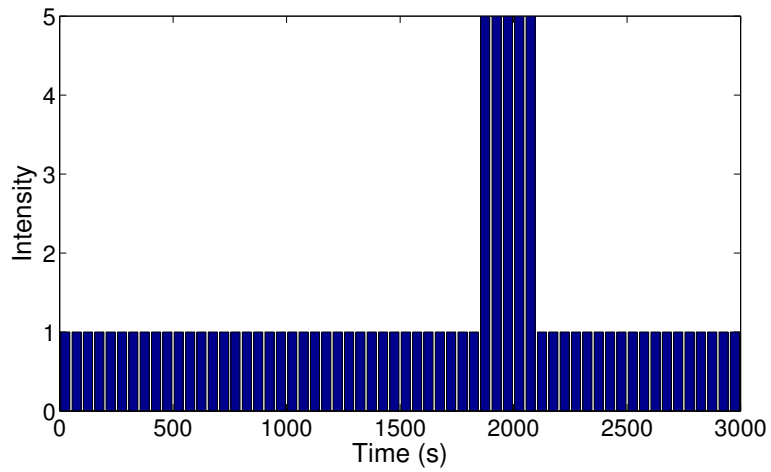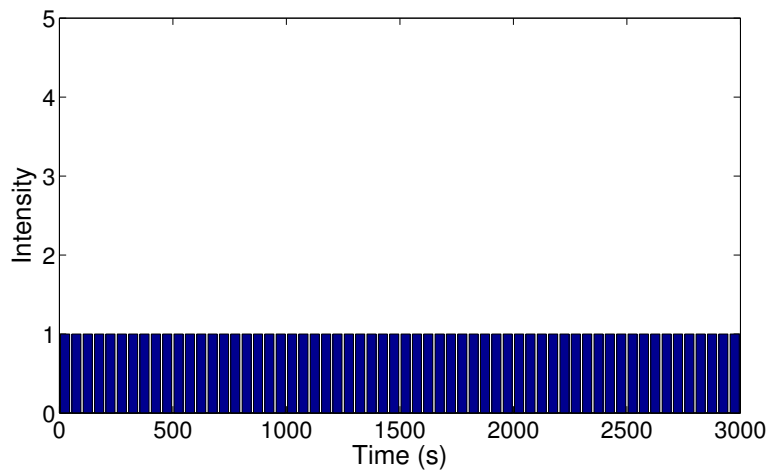
Figure 6.11: Example output from the median filter using the input signal shown in (a) using a window 3 bins wide (b) and 15 bins wide (c).

removed all of the artifacts far from the source direction in the spatial distributions seen in Figure 6.13 while preserving the sharpness in the image.

6.4 Performance Comparison

The last few sections described three different ways to improve the noisy source distributions reconstructed by the standard time-dependent MLEM algorithm. All of the methods showed significant improvement over the standard MLEM reconstruction, but which one is best? This question is difficult to answer in general; however, the ^{137}Cs data from the known-motion measurement used in Section 5.4 is a simple and fairly representative example of a situation where these algorithms may be used. Thus, in this section we find the optimal reconstruction parameters (time bin width, penalization factor, filter window) for each reconstruction for this particular measurement. Then the algorithms are directly compared using a simple test metric to determine which one performs the best.

6.4.1 Test Metric

To compare the different algorithms we must create a test metric, or figure of merit (FOM), that will quantify their performance. As discussed earlier, there are many different metrics that one might choose from to quantify performance like accuracy, resolution (FWHM), signal-to-noise ratio (SNR), and others. The resolution is difficult to calculate since these methods are based on MLEM reconstructions. The SNR is important, but for point sources perhaps it is more important that the source direction is accurate. Thus, the test metric we chose for these comparisons is the average angular distance from the true source direction to the direction of the hottest
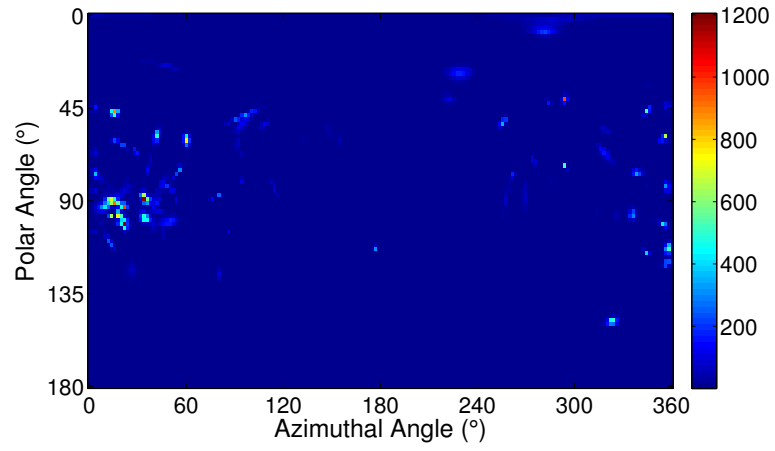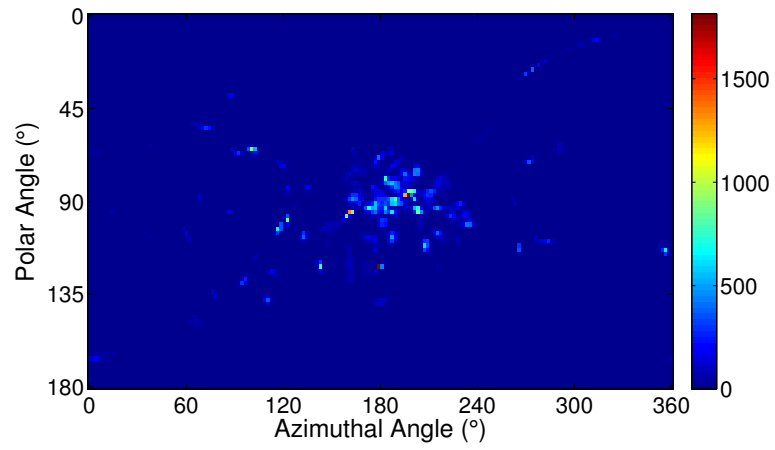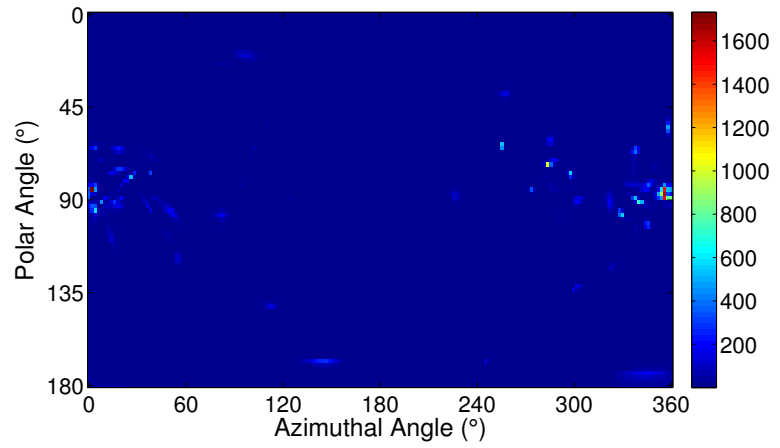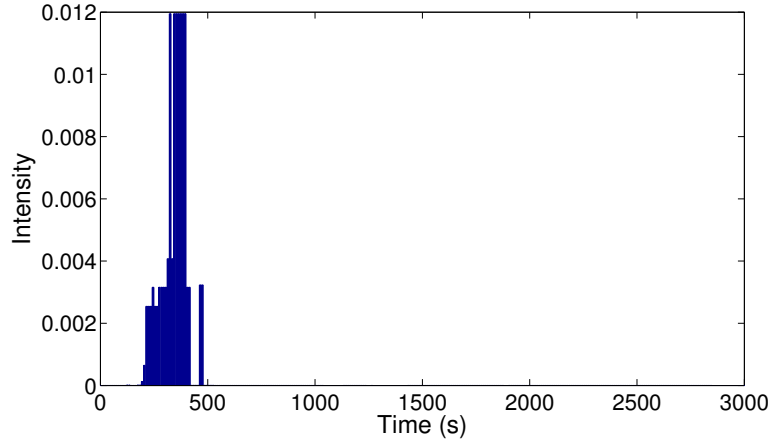
(a)
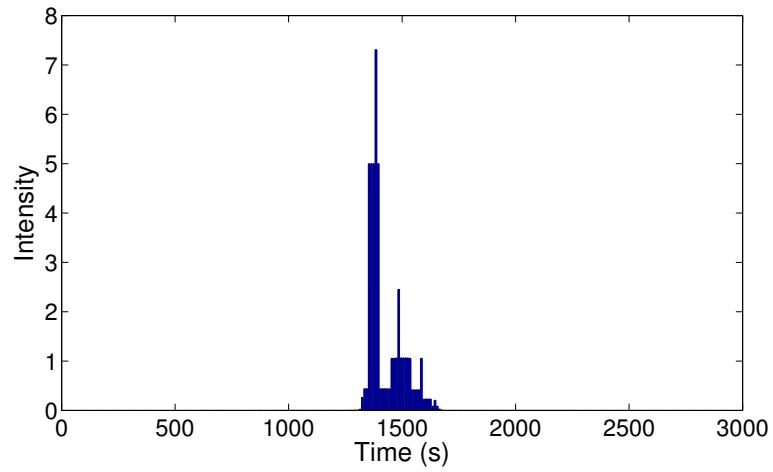
(b)
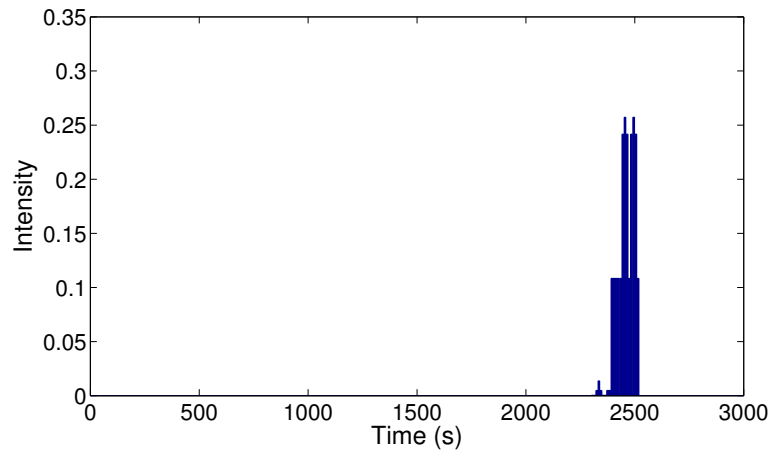
(c)

Figure 6.12: Unregularized MLEM solution after 25 EM iterations post-processed using the median filter with a window of 21 bins. The plots show the source intensity as a function of time along the equator at an azimuthal angle of (a) 45°, (b) 180°, and (c) 315°.

154
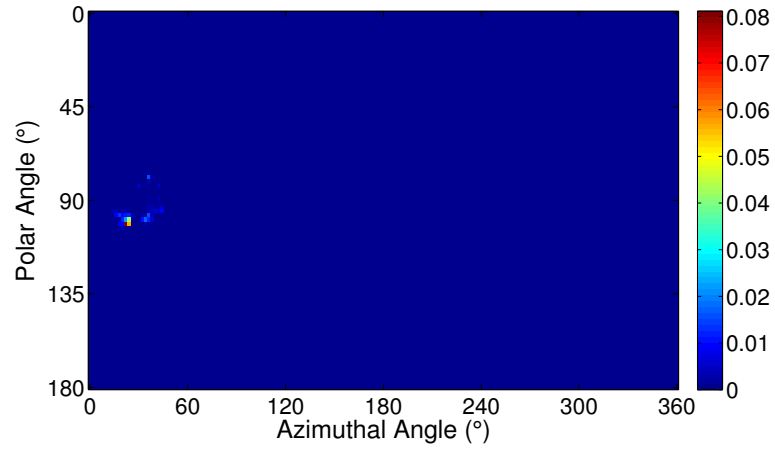
(a)
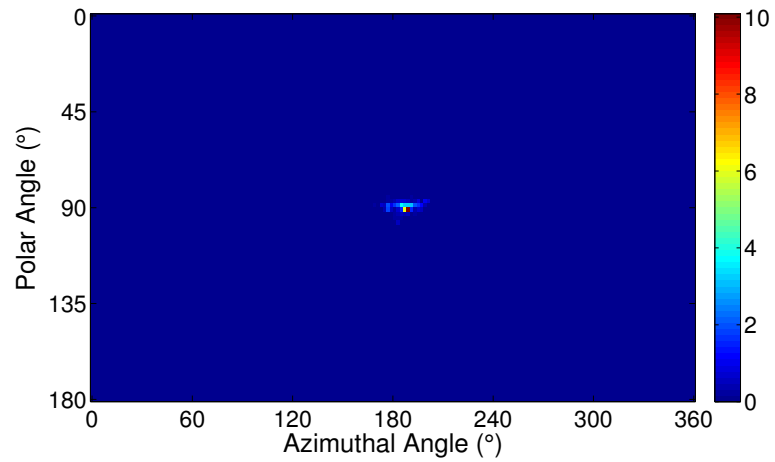
(b)
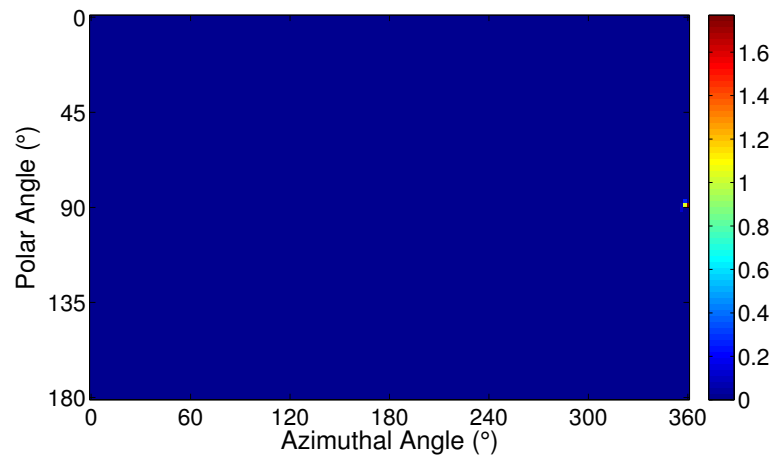
(c)

Figure 6.13: Unregularized MLEM solution after 25 EM iterations post-processed using the median filter with a window of 21 bins. The figures show the spatial source distribution at (a) 230 s, (b) 1500 s, and (c) 2800 s.

pixel in the image:

$$FOM = \frac{1}{L} \sum_{l=1}^{L} \sqrt{(\theta_l^{true} - \theta_l^{max})^2 + (\phi_l^{true} - \phi_l^{max})^2} \qquad (6.27)$$

where θ_l^{true} is the true polar angle of the source at the center of time bin l and θ_l^{max} is the polar angle corresponding to the hottest spot in the reconstructed image at time bin l. The same is true for the azimuthal directions ϕ_l^{true} and ϕ_l^{max}. Remember that the true source position is known for this dataset because the source was attached to a computer-controlled actuator arm.

We must take some care to ensure that the azimuthal angle is handled correctly near the prime meridian ($\phi = 0$). Specifically, if the true source direction is at an azimuthal angle of 5° and the reconstructed max azimuthal angle is 355°, the above formula would result in a very large FOM with an azimuthal angle offset of 350°. However, the true offset is only 10°. Thus, the azimuthal difference term in (6.27) is interpreted as

$$(\phi_l^{true} - \phi_l^{max})^2 =$$
$$\min \left[(\phi_l^{true} - \phi_l^{max})^2, (\phi_l^{true} - \phi_l^{max} - 360°)^2, (\phi_l^{true} - \phi_l^{max} + 360°)^2 \right]. \qquad (6.28)$$

6.4.2 Parameter Selection

The free parameters in this study are the number of evenly spaced time bins L, the penalty factor δ, and the filter window sizes W_{avg} and W_{med} for the average filter and median filter, respectively. Each reconstruction method used in this comparison has two free parameters, except for the standard time-dependent MLEM reconstruction which only has the number of time bins. In addition to the number of time bins, the penalized-likelihood also has the penalty factor, and the filter reconstructions have their respective filter windows. For simplicity, we fixed the number of iterations at

25 for both the standard time-dependent MLEM reconstruction and the penalized-likelihood reconstruction.

The first step of the optimization is to select the set of L values to use for the reconstructions. Since the measurement used for this comparison lasted approximately 3000 s,we used a time range of 0 s -3000 s, and since the UMImaging software requires bins widths to be integers, we used all of the factors of 3000 between 10 and 750 (*i.e.,* 10, 12, 15, 20, ... 500, 600, 750) for the number of time bins for each reconstruction. The penalty factors used for the regularized MLEM reconstruction were

$$\delta = \{0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1.0\}$$

which through experience was found to be a reasonable range of values. Finally, we increased the size of the windows used for filtering the standard MLEM reconstructions until the FOM reached a minimum for each number of time bins.

We performed each reconstruction for all combinations of parameters to find the best FOM for each reconstruction type.

6.4.3   Results

For each proposed reconstruction type, we analyze the FOM as a function of number of time bins and penalization factor (or filter window width).

6.4.3.1   Regularized Time-Dependent MLEM

Figure 6.14 shows the calculated FOM for the regularized time-dependent MLEM algorithm over the proposed parameter space. As the penalization factor δ approaches zero, the regularized solution approaches the unregularized solution. Therefore, the expected FOM for the unregularized case can be approximated by the case where $\delta = 0.001$. When using that small amount of regularization, the best FOM values are

found where there are relatively few time bins (about 30). For this scenario, there are enough counts in each time bin (about 80) to reconstruct an accurate estimate of the source direction without having too much blurring due to the wide time bins. As the number of time bins increases, the poor counting statistics cause the source direction estimate to be increasingly inaccurate. This effect would be exaggerated with even lower values of δ.

As the penalization parameter increases, the optimal number of time bins increases as well. This result should be expected since a major assumption we made was that the source motion is smooth in time. Therefore, it is desirable to increase the number of time bins which reduces the amount of fluctuation between successive bins, but more regularization (higher values of δ) is then required to reduce the influence of statistical fluctuations. There reaches a practical limit however, where the marginal gains are too small to offset the extra computation required for the increasing number of time bins. This point may be above 200 or 300 time bins in this example since the FOM essentially flattens out for reasonable penalization factors above that.

Another point to consider is that there are other factors not encapsulated by the simple test metric we are using to calculate this FOM. Although the estimate of the source position is improving with more time bins, the resolution in space and time may be fluctuating and the signal-to-noise ratio may also be changing. However, for this simple test metric the best results were obtained for 750 time bins and $\delta = 0.1$, but the bottom of the curves are quite flat; so there is a range of reasonable parameter choices.

6.4.3.2 Average Filtered Time-Dependent MLEM

Figure 6.15 shows the distribution of FOM values for the solutions calculated by applying the average filter to the time-dependent MLEM results. The biggest difference between these results and the regularized MLEM results is the FOM as
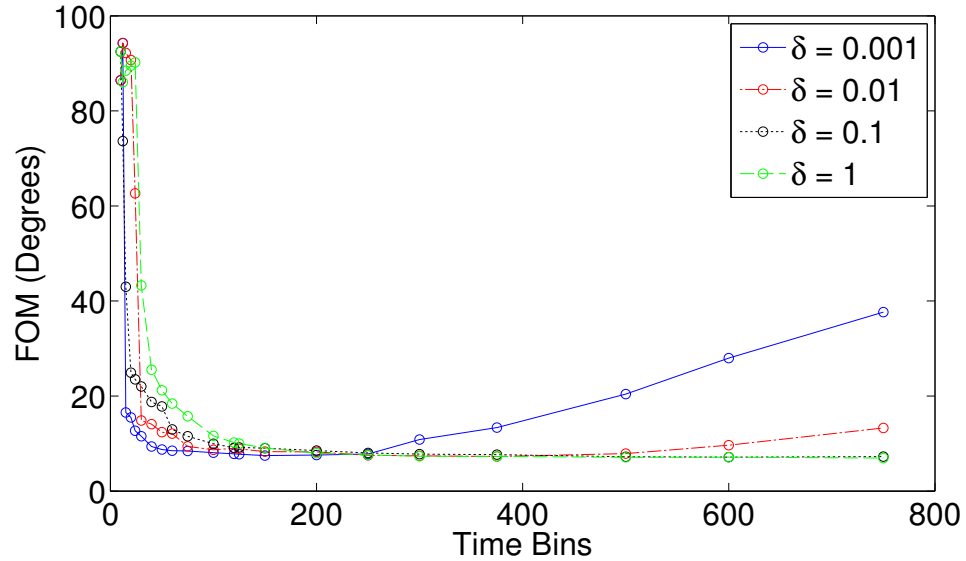
Figure 6.14: FOM results for the regularized time-dependent MLEM algorithm.

the number of time bins increases beyond 200. The average filter does not effectively remove the statistical noise from the standard MLEM reconstruction when the counts are spread over so many time bins. This observation agrees with the imaging results where the average filter struggled to reconstruct a readily identifiable hotspot for a reconstruction using 300 time bins.

Another interesting trend is that as the size of the filter window increases, the FOM for the lower numbers of time bins gets worse. This effect is a result of the filter averaging over a large absolute time which smears the reconstructed source distribution through space.

Because the filter does not effectively reduce the noise in the reconstruction, the best option is to have 100 time bins and a filter window of 7 bins. However, this is not much different than just the standard unfiltered MLEM reconstruction, and any benefits of using this filter may not be worth the extra implementation or computation time.
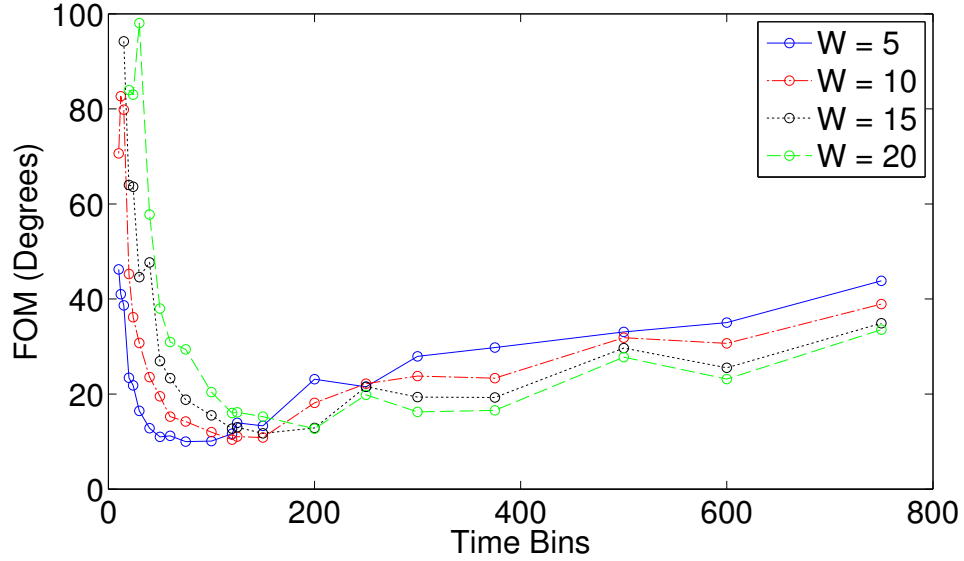
Figure 6.15: FOM results for the time-dependent MLEM solution after applying the average filter.

6.4.3.3 Median Filtered Time-Dependent MLEM

Figure 6.16 shows the FOM for the results obtained using the unregularized MLEM solution and applying the median filter. These results are a definite improvement over the results obtained from the average filter. Notice that the overall FOM values are lower and more comparable to the regularized MLEM results. Also, even though the FOM does increase with the number of time bins, it does not increase as rapidly as the average-filtered results. In fact, for larger filter widths, the minimum point on the FOM curve is actually extended out to larger numbers of time bins.

This improved performance for large numbers of time bins occurs for two reasons. First, the statistical noise is easily filtered out with the median filter, especially when many time bins are used and the artifacts becomes more randomly distributed in space and time. At the same time, if the filter window is wide enough, it can still capture the true source position if at least half of the time bins in the window have a hotspot in the source direction. However, inevitably the number of counts in each time bin will be low enough that the reconstruction cannot estimate the correct source

160
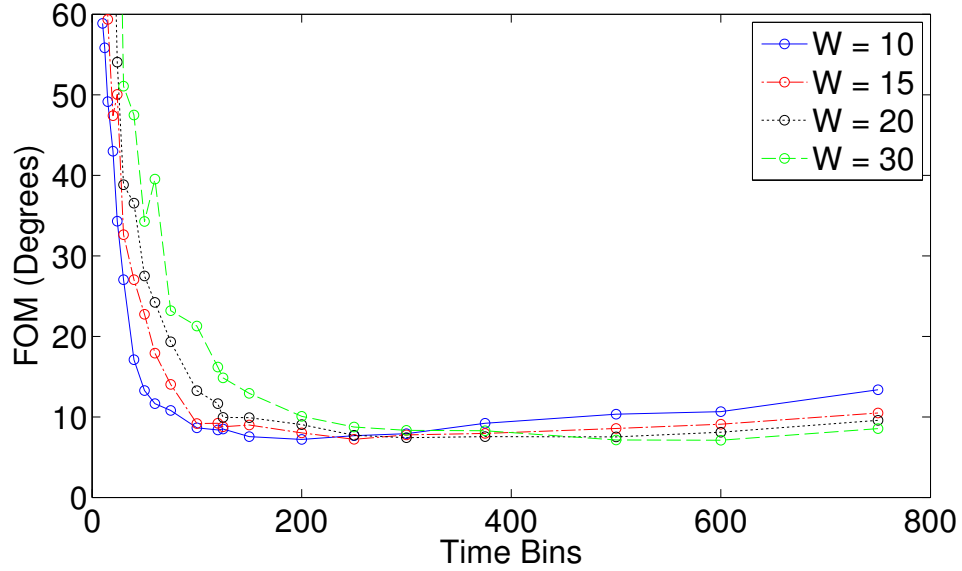
Figure 6.16: FOM results for the time-dependent MLEM solution after applying the median filter.

position. For this reconstruction the best results are obtained when 500 time bins are used with a filter window 28 bins wide.

6.5   Discussion

The work in this chapter has described several different approaches to handling time-dependent image reconstruction with a limited count rate. For the simple experiment on which we focused, the regularized time-dependent MLEM algorithm and the median-filtered time-dependent MLEM seemed to produce the best results. The parameter optimization performed in the previous section gives us some important insight into the behavior of these algorithms. Table 6.1 summarizes the best results for each reconstruction. However, as seen by the small differences between the optimal FOV values, not only does the FOM need to include other metrics than just the source position (*e.g.,* resolution or SNR), but it is also an impractical procedure to follow each time the reconstruction is performed. Further study is required to fully flesh out the selection process for the number of time bins to use as well as the penalty

Table 6.1: Optimal parameters and FOMs for time-dependent algorithms.

Reconstruction Method	Time Bins	δ / Filter Width	FOM (°)
Standard Time-Dependent MLEM	30	N/A	7.0
Regularized MLEM	750	0.1	6.9
Average-Filtered MLEM	100	7	9.5
Median-Filtered MLEM	500	28	6.8

factor or the filter window widths.

Despite the simplicity of the analysis, some guidelines for selecting the reconstruction parameters may be gleaned from the results in the previous section. First, there must be enough time bins to reduce the amount of source motion in any given time interval. For example, in the experiment used for this chapter, the source moved 360° in approximately 3000 s (0.12 degrees/s). For the regularized MLEM algorithm, the optimal number of time bins was about 750. This binning results in approximately 0.5° of source motion in a given time bin. The imaging mesh used was 90×180 bins in the polar and azimuthal directions respectively. Thus, the source direction essentially appears stationary over any 4 time bins. A good rule of thumb might be to have at least enough time bins such that for the fastest projected source speed it takes at least 2 time bins to move through a spatial bin. This way the smooth assumption in the algorithm is upheld. Similar types of arguments could be made for the other two reconstruction methods.

Selecting a penalization factor or filter windows is also a challenge and is an ongoing topic of research for many types of problems and regularization schemes [69–71]. For our case, one way to roughly estimate the necessary penalization factor may be to use the count rate. If there are many counts in each time bin, then the regularization will be less necessary and δ can be a low value. The inverse is also true; if there are few counts in each bin, more regularization is necessary and δ should be relatively high. Of course these are heuristic guidelines, but after some study a reasonable range could be determined for a typical scenario.

CHAPTER VII

Summary and Future Work

7.1 Summary and Conclusions

Compton imaging using 3-D-position-sensitive CdZnTe detectors is a powerful tool for imaging and localizing gamma-ray sources of energies above 300 keV. This imaging capability combined with excellent energy resolution make these detectors ideal for many applications in the fields of nuclear industry, medical imaging, and defense and homeland security. This work has reviewed algorithms for 2-D and 3-D image reconstruction, demonstrated the ability to locate multiple point sources in 3-D using a moving detector, and introduced two new image-reconstruction methods for use with moving sources.

The review of current 2-D imaging methods focused on practical issues and important options as well as their performance. These issues include the need for interaction-sequence reconstruction and the full-energy-deposition assumption. Ring normalization and acceptable interaction-separation distances were key reconstruction options that have the potential to drastically affect the quality of the reconstructed image. The MLEM review included several different system-model options including simple, standard, and complete. The simple model is the most computationally efficient, but for complicated geometries or source distributions, the standard system model can provide the best spatial resolution. The complete model is the most

computationally expensive, but it can provide the user with an estimate of the true incident energy spectrum as a function of direction. Although the angular resolution of SBP (30°) is poor compared to the MLEM reconstruction using the standard system model (10°), SBP takes fewer counts to achieve the same source-direction uncertainty.

Chapter IV discussed several different methods for 3-D Compton imaging. Near-field 3-D imaging is possible even with a single detector. When multiple detectors are available, the reconstruction algorithms can take advantage of the added parallax to expand the filed of view to greater distances. A simple geometrical analysis provided the user with a rough idea of the expected FOV given basic knowledge about the angular resolution and dimensions of the detector geometry. A detector in motion is a powerful tool for search scenarios since it allows the user to make the most efficient use of parallax. Experimental results using a moving detector demonstrated the feasibility of using a single imaging detector to make a 3-D radiation map of a room while providing an intuitive way to view the results by projecting them onto an optical image. However, this type of reconstruction requires that the position of the detector is known as a function of time, and as the reconstruction space increases, the memory and computational cost of reconstructing on a 3-D mesh is restrictive. Thus, more work is required to make this reconstruction practical, *e.g.,* multi-scale meshes or special caching techniques, especially if higher resolution is desired.

Moving sources introduce a new challenge, but Chapter V introduced a new reconstruction method that simplifies the problem when the source motion is known. Initially, one may perform simple motion compensation to keep the source in focus. However, standard motion compensation only allows the reconstruction to focus on a single target. To overcome this limitation, the new method adds new time-dependent spatial bins. The time-dependence of these bins compensates for the motion of any number of targets in the FOV. By including this new binning structure, the recon-

struction can estimate the source distribution for the stationary backdrop space as well as any moving targets without introducing interference between them. Experimental results show that this new reconstruction can successfully deconvolve the moving-source and stationary-source distributions. However, as with the complete model MLEM reconstruction in the static case, this reconstruction typically provides poor angular resolution and is computationally expensive. With limited statistics or computational power, or when the reconstruction of the incident spectrum has limited value, one could collapse the energy domain and only reconstruct photopeak events, *i.e.,* use the standard system model with added spatial basis functions instead of the complete system model. This form would be more efficient and potentially provide better angular resolution.

However, when there is no way to estimate the motion of potential source targets, the reconstruction must explicitly include the time dimension. This added dimension makes an already ill-posed problem worse. To reduce the effects of poor statistics, we add a regularization term to the MLEM reconstruction. This new term encourages the solution to be smooth in time and reduces the artifacts found in the unregularized solution caused by poor statistics. We also use rolling average and median filters to smooth out the time domain when applied to the unregularized-MLEM solution. The rolling average filter does not perform well, but the median filter is well suited to remove the "salt and pepper" type of noise present in the time domain, and it produces similar results to that of the regularized MLEM reconstruction. Further analysis including new test statistics, *e.g.,* SNR or angular resolution, is required to give a more complete comparison between these methods. However, it is obvious that these reconstruction techniques can provide much improved images of moving sources especially for count-limited measurements. Also, a simple or automated way to choose parameters like the penalization factor δ and the filter window sizes must be implemented to make the reconstructions practical for end users.

7.2 Future Work

The discussion sections of the last few chapters included some ideas for improvement or new applications for the reconstruction methods presented in this work. This final section will describe other important improvements that are necessary to take these reconstructions to the next level and become more practical for use in the field.

7.2.1 Compton Cone Reconstruction

In the current implementation, the angle between the axis of the Compton cone and the center of each imaging pixel determines if that pixel falls within the uncertainty of the Compton cone projection. This method is straightforward to implement, but makes the reconstruction time linearly dependent on the number of imaging pixels, which directly conflicts with the desire to increase the number of pixels to improve spatial resolution. Thus, we need a new way to reconstruct the Compton cones that minimizes the number of required calculations or find a way to speed up those calculations.

One way to reduce the number of calculations is to use a marching algorithm like the one described by Wilderman et al. [72]. Instead of blindly testing every pixel in the image to determine if it lies on the Compton cone, one can imagine initially searching for a point on the ring, then using the geometry of the intersection of a cone with the surface to only perform calculations where the ring intersects the pixels. Although this was originally implemented for 3-D, it could be adapted to the 2-D case as well.

Another way to speed up the calculation time is to avoid using a mesh for the image reconstruction [73,74]. In this origin ensembles method, instead of reconstructing the entire Compton cone on a 3-D mesh, a single point (or *origin*) is randomly placed at a position that lies on the Compton cone. After this procedure is repeated for all events, the reconstructed points are moved stochastically on their conical surfaces, and the acceptance of the new positions is based on a predefined acceptance probability

defined by the new point density of the region. After many repetitions of this process, the solution converges to the source distribution estimate.

A similar type of reconstruction was recently presented by Haefner et al. [75]. In this variant, three events are chosen at random, and the position that is closest to the intersection of these cones is set as a reconstructed point in the image. This process is repeated many times to form a point cloud image. The events used for a given repetition are not removed indefinitely but are placed back into the pool of events. This method and the origin ensembles method would need extra care to ensure that partial energy deposition events are handled properly.

Finally, instead of completely altering the reconstruction method, one could more efficiently use the power of modern computers. Advances in graphical processing units (GPUs) have made it possible to perform massively parallel calculations much faster than on standard CPUs. Since the system response calculations are readily parallelized, this kind of implementation has the potential to reduce the calculation time by an order of magnitude or more [44].

7.2.2   Position and Orientation Estimation

The development of a truly handheld 3-D radiation imaging device would require a robust system to determine its position and orientation. This problem is a current research topic with a lot of attention in the autonomous-robotics field [76–79]. These techniques usually employ accelerometer based technology, standard optical cameras, or depth-sensing optical cameras. Accelerometers are inertial devices that estimate position and orientation based on the integration of the recorded acceleration data. Alternatively, the camera solutions use reference points in the optical image to estimate the orientation and position of the system. A bonus of the camera systems is that they can employ simultaneous localization and mapping (SLAM) algorithms that allow the device to know where it is and make a map of the space it has tra-

versed. The camera systems are also attractive because cameras are already a part of the current detector system.

7.2.3 Beyond Imaging

This work has focused on Compton imaging. However, even though the source-distribution estimate is useful for many scenarios, sometimes other information is more important.

7.2.3.1 Source Characterization

Once a source has been found, the user may want to know more about it than just its position and the energy of its emissions. This interesting analysis will generally happen relatively close to the source, and most sources have some extent when viewed in the near field. Not only can the shape of the source object be determined, but also a well-calibrated detector can begin to characterize the source in terms of its mass and isotopic composition. This type of analysis goes beyond the simple image reconstruction and requires new algorithms to estimate these parameters. However, even though the desired information is not the source distribution, the new algorithms may take advantage of the imaging capability. For example, the accuracy of estimate of the mass may be improved if the reconstructed image could determine the shape and volume of the source.

7.2.3.2 Detection

Source detection and identification is another important application. Determining if a source is present in a given measurement is more difficult than it may initially seem. Statistical uncertainty and fluctuations in background intensities make the task more challenging. Previous work by Wahl [80] showed excellent detection performance in an unknown background environment by taking advantage of the spectroscopic

and imaging capabilities of pixelated CdZnTe. Many of the reconstruction methods described in this dissertation could be easily adapted to work with the detection methods presented by Wahl. Specifically, the method that takes advantage of known source motion would integrate almost seamlessly, which would allow the detection algorithm to work with moving sources.

However, the detection algorithms are slow and only work for energies above 300 keV where Compton imaging is possible. These traits are undesirable because detection is usually time sensitive and many sources of interest emit gamma rays with energies below 300 keV most of the time. Thus, improvements must be made that allow for quicker detection times and the ability to detect and identify sources emitting low-energy gamma rays.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Inductiveload and NASA, "Em_spectrum_properties_edit.png." http://en.wikipedia.org/wiki/File:EM_Spectrum_Properties_edit.svg, Oct. 2007. Accessed Dec. 2012.

[2] M. Berger, J. Hubbell, S. Seltzer, J. Chang, J. Coursey, R. Sukumar, D. Zucker, and K. Olsen, "XCOM: Photon cross sections database." http://www.nist.gov/pml/data/xcom/index.cfm, Dec. 2011.

[3] G. F. Knoll, *Radiation Detection and Measurement.* John Wiley & Sons, 2010.

[4] N. Tsoulfanidis, *Measurement and Detection of Radiation.* Taylor & Francis, 1995.

[5] D. Xu, *Gamma-Ray Imaging and Polarization Measure Using 3-D Position-Sensitive CdZnTe Detectors.* PhD thesis, University of Michigan, 2006.

[6] W. Wang, *Techniques and Applications of Compton Imaging for Position-Sensitive Gamma-Ray Detectors.* PhD thesis, University of Michigan, 2011.

[7] C. G. Wahl, *Imaging, Detection, and Identification Algorithms for Position-Sensitive Gamma-Ray Detectors.* PhD thesis, University of Michigan, 2011.

[8] C. G. Wahl, J. M. Jaworski, and Z. He, "UMImaging: A software package for image reconstruction from 3D-position-sensitive gamma-ray detectors," *IEEE Trans. Nucl. Sci.*, vol. 59, pp. 1672 –1680, Aug. 2012.

[9] W. Shockley, "Currents to conductors induced by a moving point charge," *Journal of Applied Physics*, vol. 9, pp. 635 –636, Oct. 1938.

[10] S. Ramo, "Currents induced by electron motion," *Proceedings of the IRE*, vol. 27, no. 9, pp. 584 – 585, 1939.

[11] Z. He, "Review of the Shockley–Ramo theorem and its application in semiconductor gamma-ray detectors," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 463, no. 12, pp. 250 – 267, 2001.

[12] P. N. Luke, "Single-polarity charge sensing in ionization detectors using coplanar electrodes," *Applied Physics Letters*, vol. 65, no. 22, pp. 2884–2886, 1994.

[13] D. McGregor, R. Rojeski, Z. He, D. Wehe, M. Driver, and M. Blakely, "Geometrically weighted semiconductor Frisch grid radiation spectrometers," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 422, no. 13, pp. 164 – 168, 1999.

[14] V. Ivanov, L. Aleksejeva, P. Dorogov, and A. Luchanskii, "CZT quasi-hemispherical detectors with improved spectrometric characteristics," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pp. 1696 –1699, Nov. 2009.

[15] Z. He, W. Li, G. Knoll, D. Wehe, J. Berry, and C. Stahle, "3-D position sensitive CdZnTe gamma-ray spectrometers," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 422, pp. 173–78, 1999.

[16] H. H. Barrett, J. D. Eskin, and H. B. Barber, "Charge transport in arrays of semiconductor gamma-ray detectors," *Phys. Rev. Lett.*, vol. 75, pp. 156–159, July 1995.

[17] Z. He, G. F. Knoll, D. K. Wehe, R. Rojeski, C. H. Mastrangelo, M. Hammig, C. Barrett, and A. Uritani, "1-D position sensitive single carrier semiconductor detectors," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 380, no. 12, pp. 228 – 231, 1996.

[18] F. Zhang, Z. He, D. Xu, G. Knoll, D. Wehe, and J. Berry, "Improved resolution for 3-D position sensitive CdZnTe spectrometers," *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 2427 – 2431, Oct. 2004.

[19] S. Kaye, W. Kaye, J. Jaworski, and Z. He, "Experimental limitations of coded aperture imaging using thick 3D-position-sensitive CdZnTe detectors," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE*, pp. 3856 –3859, Nov. 2010.

[20] J. C. Kim, W. R. Kaye, W. Wang, F. Zhang, and Z. He, "Impact of drift time variation on the Compton image from large-volume CdZnTe crystals," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 683, no. 0, pp. 53 – 62, 2012.

[21] F. Zhang and Z. He, "New readout electronics for 3-D position sensitive CdZnTe/HgI$_2$ detector arrays," *IEEE Trans. Nucl. Sci.*, vol. 53, pp. 3021 –3027, Oct. 2006.

[22] W. Wang, W. R. Kaye, J. C. Kim, F. Zhang, and Z. He, "Improvement of Compton imaging efficiency by using side-neighbor events," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 687, no. 0, pp. 62 – 68, 2012.

[23] H. Yang, J. C. Kim, and Z. He, "Event classification and reconstruction of single pixel multiple interaction events," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pp. 4663 –4667, Oct. 2011.

[24] S. A. et al., "Geant4–a simulation toolkit," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 506, no. 3, pp. 250 – 303, 2003.

[25] D. Habs, M. M. Günther, M. Jentschel, and W. Urban, "Refractive index of silicon at γ ray energies," *Phys. Rev. Lett.*, vol. 108, p. 184802, May 2012.

[26] M. Singh, "An electronically collimated gamma camera for single photon emission computed tomography. Part I: Theoretical considerations and design criteria," *Medical Physics*, vol. 10, no. 4, pp. 421–427, 1983.

[27] M. Singh and D. Doria, "An electronically collimated gamma camera for single photon emission computed tomography. Part II: Image reconstruction and preliminary experimental measurements," *Medical Physics*, vol. 10, no. 4, pp. 428–435, 1983.

[28] T. Nagayoshi, H. Kubo, K. Miuchi, A. Ochi, R. Orito, A. Takada, T. Tanimori, and M. Ueno, "Performance of large area micro pixel chamber," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 513, no. 12, pp. 277 – 281, 2003.

[29] D. Chivers, A. Coffer, B. Plimley, and K. Vetter, "Impact of measuring electron tracks in high-resolution scientific charge-coupled devices within Compton imaging systems," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 654, no. 1, pp. 244 – 249, 2011.

[30] D. Xu, Z. He, C. E. Lehner, and F. Zhang, "4π Compton imaging with single 3-D position sensitive CdZnTe detector," in *Proc. of SPIE*, vol. 5540, SPIE, 2004.

[31] C. L. Thrall, C. G. Wahl, and Z. He, "Performance of five-or-more-pixel event sequence reconstruction for 3-D semiconductor gamma-ray-imaging spectrometers," in *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pp. 1299 –1301, Oct. 2008.

[32] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *Medical Imaging, IEEE Transactions on*, vol. 1, pp. 113 –122, Oct. 1982.

[33] T. Hebert, R. Leahy, and M. Singh, "Three-dimensional maximum-likelihood reconstruction for an electronically collimated single-photon-emission imaging system," *J. Opt. Soc. Am. A.*, vol. 7, pp. 1305–13, 1990.

[34] S. Wilderman, N. Clinthorne, J. Fessler, and W. Rogers, "List-mode maximum likelihood reconstruction of Compton scatter camera images in nuclear medicine," in *Nuclear Science Symposium, 1998. Conference Record. 1998 IEEE*, vol. 3, pp. 1716 –1720 vol.3, 1998.

[35] H. H. Barrett, T. White, and L. C. Parra, "List-mode likelihood," *J. Opt. Soc. Am. A*, vol. 14, pp. 2914–2923, Nov. 1997.

[36] L. C. Parra and H. H. Barrett, "List-mode likelihood: EM algorithm and image quality estimation demonstrated on 2-D PET," *IEEE Trans. Med. Imag.*, vol. 17, pp. 228–235, Apr. 1998.

[37] F. Zhang, *Events reconstruction in 3-D position sensitive CdZnTe gamma-ray spectrometers.* PhD thesis, University of Michigan, 2005.

[38] W. R. Kaye, *Energy and Position Reconstruction in Pixelated CdZnTe Detectors.* PhD thesis, University of Michigan, 2012.

[39] D. Xu and Z. He, "Gamma-ray energy-imaging integrated spectral deconvolution," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 574, pp. 98–109, Apr. 2007.

[40] W. Wang, C. Wahl, J. Jaworski, and Z. He, "Maximum-likelihood deconvolution in the spatial and spatial-energy domain for events with any number of interactions," *IEEE Trans. Nucl. Sci.*, vol. 59, pp. 469–478, Apr. 2012.

[41] D. J. Lingenfelter and J. A. Fessler, "System modeling for gamma-ray imaging systems," Tech. Rep. 411, University of Michigan, Mar. 2012.

[42] D. Nagirner and V. Loskutov, "Compton attenuation coefficient in scattering by Maxwellian electrons," *Astrophysics*, vol. 43, pp. 343–350, 2000. 10.1007/BF02683970.

[43] A. Sauve, I. Hero, A.O., W. Rogers, S. Wilderman, and N. Clinthorne, "3D image reconstruction for a Compton SPECT camera model," *IEEE Trans. Nucl. Sci.*, vol. 46, pp. 2075 –2084, Dec. 1999.

[44] V.-G. Nguyen, S.-J. Lee, and M. N. Lee, "GPU–accelerated 3D Bayesian image reconstruction from Compton scattered data," *Phys. Med. Biol.*, vol. 56, p. 28172836, 2011.

[45] M. W. Rawool-Sullivan, J. P. Sullivan, S. R. Tornga, and S. P. Brumby, "A simple algorithm for estimation of source-to-detector distance in Compton imaging," *Applied Radiation and Isotopes*, vol. 66, no. 12, pp. 1986–991, 2008.

[46] L. Mihailescu, K. Vetter, and D. Chivers, "Standoff 3D gamma-ray imaging," *IEEE Trans. Nucl. Sci.*, vol. 56, pp. 479–486, Apr. 2009.

[47] J. R. Clynch, "GPS accuracy levels." http://www.oc.nps.edu/oc2902w/gps/gpsacc.html, Feb. 2001. Accessed Nov. 2012.

[48] F. Roellinghoff, M.-H. Richard, M. Chevallier, J. Constanzo, D. Dauvergne, N. Freud, P. Henriquet, F. L. Foulher, J. Ltang, G. Montarou, C. Ray, E. Testa, M. Testa, and A. Walenta, "Design of a Compton camera for 3D prompt-imaging during ion beam therapy," *Nucl. Inst. Meth. Phys. Res. A.*, vol. 648, Supplement 1, no. 0, pp. S20 – S23, 2011.

[49] M. Chandy, C. Pilotto, and R. McLean, "Networked sensing systems for detecting people carrying radioactive material," in *Networked Sensing Systems, 2008. INSS 2008. 5th International Conference on*, pp. 148 –155, June 2008.

[50] E. Asma, T. Nichols, J. Qi, and R. Leahy, "4D PET image reconstruction from list mode data," in *Nuclear Science Symposium Conference Record, 2000 IEEE*, vol. 2, pp. 15/57–15/65, 2000.

[51] A. J. Reader, F. C. Sureau, C. Comtat, R. Trbossen, and I. Buvat, "Joint estimation of dynamic PET images and temporal basis functions using fully 4D ML-EM," *Physics in Medicine and Biology*, vol. 51, no. 21, pp. 5455–5474, 2006.

[52] X. Niu, Y. Yang, and M. Wernick, "4D reconstruction of cardiac images using temporal Fourier basis functions," in *15th IEEE International Conference on Image Processing, 2008*, pp. 2944 –2947, Oct. 2008.

[53] T. Nichols, J. Qi, E. Asma, and R. Leahy, "Spatiotemporal reconstruction of list-mode PET data," *IEEE Trans. Med. Imag.*, vol. 21, pp. 396–404, Apr. 2002.

[54] P. M. Bloomfield, T. J. Spinks, J. Reed, L. Schnorr, A. M. Westrip, L. Livieratos, R. Fulton, and T. Jones, "The design and implementation of a motion correction scheme for neurological PET," *Physics in Medicine and Biology*, vol. 48, no. 8, p. 959, 2003.

[55] A. Rahmim, P. Bloomfield, S. Houle, M. Lenox, C. Michel, K. Buckley, T. Ruth, and V. Sossi, "Motion compensation in histogram-mode and list-mode EM reconstructions: Beyond the event-driven approach," *IEEE Trans. Nucl. Sci.*, vol. 51, pp. 2588–2596, Oct. 2004.

[56] T. Aucott, F. Jensen, and M. Bahr, "Machine vision radiation detection system," ARI, 2009.

[57] D. S. Hochbaum, "The multi-sensor nuclear threat detection problem," in *Operations Research and Cyber-Infrastructure*, vol. 47 of *Operations Research/Computer Science Interfaces Series*, pp. 389–399, Springer US, 2009.

[58] B. D. Yanoff *et al.*, "GE intelligent personal radiation locator system," ARI, 2009.

[59] Y. Censor, "Finite series-expansion reconstruction methods," *Proceedings of the IEEE*, vol. 71, pp. 409–419, Mar. 1983.

[60] J. Fessler, "Image reconstruction: Algorithms and analysis." draft of book in preparation, 2009.

[61] S. M. Ross, *Stochastic Processes*. New York: Wiley, 1983.

[62] A. De Pierro, "On the relation between the ISRA and the EM algorithm for positron emission tomography," *IEEE Trans. Med. Imag.*, vol. 12, pp. 328–333, June 1993.

[63] D. Lingenfelter, *Source Detection and Image Reconstruction with Position-Sensitive Gamma-Ray Detectors*. PhD thesis, University of Michigan, 2012.

[64] D. S. Lalush and B. M. W. Tsui, "Block-iterative techniques for fast 4D reconstruction using a priori motion models in gated cardiac SPECT," *Phys. Med. Biol.*, vol. 43, pp. 875–886, Apr. 1998.

[65] P. J. Green, "On use of the EM for penalized likelihood estimation," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 52, no. 3, pp. 443–452, 1990.

[66] T. Hebert and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Trans. Med. Imag.*, vol. 8, pp. 194 –202, June 1989.

[67] J. Fessler, E. Ficaro, N. Clinthorne, and K. Lange, "Grouped-coordinate ascent algorithms for penalized-likelihood transmission image reconstruction," *IEEE Trans. Med. Imag.*, vol. 16, pp. 166 –175, Apr. 1997.

[68] M. Jacobson and J. Fessler, "An expanded theoretical treatment of iteration-dependent majorize-minimize algorithms," *IEEE Trans. Med. Imag.*, vol. 16, pp. 2411 –2422, Oct. 2007.

[69] N. Galatsanos and A. Katsaggelos, "Methods for choosing the regularization parameter and estimating the noise variance in image restoration and their relation," *Image Processing, IEEE Transactions on*, vol. 1, pp. 322 –336, July 1992.

[70] Y. Zhang, R. Li, and C.-L. Tsai, "Regularization parameter selections via generalized information criterion," *Journal of the American Statistical Association*, vol. 105, no. 489, pp. 312–323, 2010. PMID: 20676354.

[71] S. Ramani, Z. Liu, J. Rosen, J. Nielsen, and J. Fessler, "Regularization parameter selection for nonlinear iterative image restoration and MRI reconstruction using GCV and SURE-based methods," *Image Processing, IEEE Transactions on*, vol. 21, pp. 3659 –3672, Aug. 2012.

[72] S. Wilderman, W. Rogers, G. Knoll, and J. Engdahl, "Fast algorithm for list mode back-projection of Compton scatter camera data," *IEEE Trans. Nucl. Sci.*, vol. 45, pp. 957 –962, June 1998.

[73] A. Sitek, "Representation of photon limited data in emission tomography using origin ensembles," *Phys. Med. Biol.*, vol. 53, no. 12, p. 32013216, 2008.

[74] A. Andreyev, A. Sitek, and A. Celler, "Stochastic image reconstruction method for Compton camera," in *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pp. 2985 –2988, Nov. 2009.

[75] A. Haefner, D. Gunter, L. Mihailescu, J. Maltz, and K. Vetter, "3-D Compton image reconstruction with a gridless point cloud," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2012 IEEE*, 2012.

[76] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robot. Autom. Mag.*, vol. 13, pp. 99–110, June 2006.

[77] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB–D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments," *The International Journal of Robotics Research*, vol. 31, no. 5, pp. 647–663, 2012.

[78] H.-K. Lee, K. Choi, J. Park, and H. Myung, "Self-calibration of gyro using monocular SLAM for an indoor mobile robot," *International Journal of Control, Automation and Systems*, vol. 10, pp. 558–566, 2012.

[79] B. Barshan and H. Durrant-Whyte, "Inertial navigation systems for mobile robots," *IEEE Trans. Robot. Autom.*, vol. 11, pp. 328–342, June 1995.

[80] C. Wahl and Z. He, "Gamma-ray point-source detection in unknown background using 3D-position-sensitive semiconductor detectors," *IEEE Trans. Nucl. Sci.*, vol. 58, pp. 605 –613, June 2011.