INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9408515

Collimatorless coincidence imaging

Saffer, Janet Susan Reddin, Ph.D.

The University of Arizona, 1993

Copyright ©1993 by Saffer, Janet Susan Reddin. All rights reserved.

U·M·I

300 N. Zeeb Rd.
Ann Arbor, MI 48106

COLLIMATORLESS COINCIDENCE IMAGING

by

Janet Susan Reddin Saffer

Copyright© Janet Susan Reddin Saffer 1993

A Dissertation Submitted to the Faculty of the
COMMITTEE ON OPTICAL SCIENCES (GRADUATE)

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 3

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Janet Susan Reddin Saffer entitled COLLIMATORLESS COINCIDENCE IMAGING

H. H. Barrett 7/28/93
Dissertation Director Harrison H. Barrett 3/17/92 WCB
Date

STATEMENT BY AUTHOR

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

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

SIGNED: Janet R Saffler

ACKNOWLEDGEMENTS

I initially planned to separate my acknowledgements into those who helped me with the science and those who helped me personally, but the boundaries are too blurred to so categorize them. I have been blessed to receive help of both kinds from many people. The most important person was, and continues to be, my husband Rex. His love, encouragement and example illuminate my life.

I have had the good fortune to have two advisors. Thanks go to Rick Shoemaker for supporting me through my Masters degree and serving on my dissertation committee. Special thanks to Harry Barrett for accepting me in his group, for finding time in his extremely busy schedule to instruct me, and for being a fair and honest evaluator of my work. Thanks also to the third member of my committee, Brad Barber, especially for his help in the lab.

I'd also like to thank my fellow students in the Probes Group, Michel Rogulski and Neal Hartsough, for their laughter and advice. Thanks to the rest of Harry's students whose many ideas made the research environment exciting, including Ted Gooley, Sylvia Rogers, Tim Roney and my officemate, Tim White. Particular thanks to John Aarsvold for math help and for the myriad tasks he performed to help the group even at the expense of his own research. Thanks also to Jim Woolfenden, Ted Bowen and Jack Denny for discussions and helpful advice.

Special thanks to the women scientists who are my friends, especially my "little sis" Liz Alvarez and Connie Walker, the sweetest person I know and one of the strongest. Thanks also to Kelly Rehm and Heidi Schlitt for sympathy and Winnie the Pooh, to Grace Wolf and to Kim Dow.

Finally, I would like to thank my parents for teaching me that my opinion was worth hearing and that I could be anything I had the courage to aspire to and work toward. They both have been my role models and I am greatly saddened that my father did not live to see me receive my degree.

This work was supported by NCI grants PO1 CA23417 and RO1 CA52643.

TABLE OF CONTENTS

	Page
LIST OF ILLUSTRATIONS	7
LIST OF TABLES	10
ABSTRACT	11
1. INTRODUCTION	13
Scope of this Dissertation	17
2. COINCIDENCE IMAGING	20
Definition of coincidence imaging	20
Earlier work in coincidence imaging	22
Derivation of H matrix expression	32
True versus accidental coincidences	36
3. MONTE CARLO SIMULATIONS	41
Description of probe geometry	41
Step-by-step procedure used in Monte Carlo simulations	42
Test of the random number generator	44
Accidental coincidences	46
Object models used -- geometry and activities	48
Collimator comparisons	54
4. RECONSTRUCTION OF OBJECT FROM DATA	56
Sizes of f , g and H	56
Inverting the system matrix H	57
Singular-value decomposition of $H^t H$	58
Discussion of the eigenvectors of $H^t H$	61
Reconstruction of object distributions	74
Point source results	76
"Spheres-in-a-cube" results	79
Zubal phantom results	82
5. SEMICONDUCTOR DETECTOR PHYSICS	85
Semiconductor materials	85
The motion of electron-hole pairs	88
Drift velocity	90
Transit time of electrons and holes in the detector	91

TABLE OF CONTENTS - continued

	Page
The change in voltage at the detector electrodes	91
Trapping effects	93
Pulse height spectrum	95
Windowing the pulse height spectrum	99
Real world detectors versus ideal detectors	99
 6. LABORATORY EXPERIMENTS	 101
Laboratory setup	103
Determining that the circuit measures true coincidences	106
Deciding how much to lengthen the TTL pulse	109
Placement of detectors and decription of grid coordinate system	112
Determining the coordinates of the detectors -- scans in x	112
Scans in z	113
Data reduction	115
The variation in accidental and true coincidence rates with activity	119
Effect of an attenuating medium	122
Results	125
 7. SUMMARY AND SUGGESTIONS FOR FURTHER RESEARCH.	 126
Summary.	126
Suggestions for future research.	127
 REFERENCES	 130

LIST OF ILLUSTRATIONS

Figure		Page
1-1	The Anger Camera	14
1-2	Cutaway View of the Dual Probe	16
1-3	21-Element Imaging Probe	18
2-1	Common Field-Of-View of Two Collimated Detectors	21
2-2	Level Diagram of ^{111}In	23
2-3	Level Diagram of ^{123}I	24
2-4	Schmitz-Feuerhake Experiment	26
2-5	Hart Experiment	28
2-6	Monahan Experiment	29
2-7	Powell Time-of-Flight Geometry	31
2-8	Two-Dimensional Detector Array Viewing a Three- Dimensional Distribution Of Activity	33
3-1	Correlation Between ^{111}In Photons	43
3-2	Two-Dimensional Display of Four-Dimensional Data Set	45
3-3	"Spheres-in-a-Cube" Object Model	49
3-4	Slice from the Zubal Torso Phantom	51
4-1	The Eigenvalue Spectrum for $\mathbf{H}^t \mathbf{H}$	60
4-2	The 10 Eigenvectors with the largest Eigenvalues, $\mathbf{u}_1$ through $\mathbf{u}_{10}$	62
4-3	Eigenvectors $\mathbf{u}_{11}$ through $\mathbf{u}_{20}$	63
4-4	Eigenvectors $\mathbf{u}_{21}$ through $\mathbf{u}_{30}$	64
4-5	Eigenvectors $\mathbf{u}_{31}$ through $\mathbf{u}_{40}$	65
4-6	Eigenvectors $\mathbf{u}_{41}$ through $\mathbf{u}_{50}$	66

LIST OF ILLUSTRATIONS - continued

Figure		Page
4-7	Eigenvectors $\mathbf{u}_{51}$ through $\mathbf{u}_{60}$	67
4-8	Eigenvectors $\mathbf{u}_{61}$ through $\mathbf{u}_{70}$	68
4-9	Eigenvectors $\mathbf{u}_{71}$ through $\mathbf{u}_{80}$	69
4-10	Eigenvectors $\mathbf{u}_{81}$ through $\mathbf{u}_{90}$	70
4-11	Eigenvectors $\mathbf{u}_{91}$ through $\mathbf{u}_{100}$	71
4-12	Eigenvectors $\mathbf{u}_{101}$ through $\mathbf{u}_{110}$	72
4-13	SVD Reconstruction of a Point Source Centered under the Detector Plane at $z = 0.05$ cm	77
4-14	SVD Reconstruction of a Point Source Centered under the Detector Plane at $z = 0.5$ cm	78
4-15	Sum of Pixels within each x-y plane in Reconstructions of Point Sources .	80
4-16	Reconstruction of the Object Distribution in Figure 3-3	81
4-17	Zubal Phantom Results	83
5-1	Band Structure of Semiconductors	86
5-2	Detector Voltage as a Function of Time	96
5-3	Typical ^{111}In Pulse Height Spectrum	97
6-1	Closeup Photograph of Detector in Laboratory Setup	102
6-2	Diagram of Coincidence Circuit	104
6-3	Photograph of Laboratory Setup	107
6-4	Graph of Counts in Channels A1 and B1 during x scan	114
6-5	Graph of Total Coincidence Count Rate as a Function of $z^2/(D_A^3 D_B^3)$.	118
6-6	Graph of True Coincidence Count Rate as a Function of $z^2/(D_A^3 D_B^3)$. .	120
6-7	Plot of Slopes of Best-Fit Lines from the Total Coincidence Data Sets vs. Activity without Distance Correction	121

LIST OF ILLUSTRATIONS - continued

Figure		Page
6-8	Plot of Slopes of Best-Fit Lines from the Total Coincidence Data Sets vs. Activity with Distance Correction	123
6-9	Graph of True and Total Coincidence Count Rates for Data Runs with and without Attenuator	124
7-1	Alternate Geometries for Coincidence Imaging	128

LIST OF TABLES

Table	Page
3-1 Organs Identified in Zubal Phantom	52
3-2 Zubal Phantom Organs Grouped by Activity	53
6-1 Effect of Varying Length of Coincidence Interval	111
6-2 List of Data Sets Used in Data Analysis	116

ABSTRACT

This dissertation explores a novel design for a surgical probe, a collimatorless coincidence imaging system designed to aid in tumor detection in nuclear medicine. Surgical probes can be maneuvered close to a suspected tumor site, thereby achieving higher resolution and sensitivity than external gamma cameras. However, conventional probes cannot distinguish between distant background variations and small tumors near the probe.

Collimatorless coincidence imaging is a new method for suppressing the effects of variations in the background radiation. This decidedly unconventional imaging system images without a collimator or aperture of any kind. The probe design consists of a 10x10 array of collimatorless gamma-ray detectors connected by coincidence circuitry. The probe is used with a radionuclide that emits multiple photons per decay, such as ^{111}In . The coincidence circuitry triggers data collection only when two photons strike the detectors within a short time interval. Because the photons are emitted independently, the probability of coincident hits on two detectors is proportional to the product of the solid angles subtended by the two detectors. Therefore distant sources have a very low probability of contributing to the data, making them all but invisible to the probe.

Data collection from such a system was simulated using a Monte Carlo routine that included absorption, the slight correlation between the directions of the emitted photons, and the presence of accidental coincidences. The data were reconstructed into object representations using the pseudoinverse obtained by singular value decomposition (SVD). The images showed a significant suppression of distant sources when compared to a probe equipped with a conventional parallel-hole collimator.

We confirmed in the laboratory, using a point source of In-111 and two CdTe detectors connected by an AND gate, that the falloff in sensitivity was inversely proportional

to the fourth power of the distance to the source and that the proportion of true to accidental coincidences followed the predicted relationship to the source activity.

We conclude that collimatorless coincidence imaging is promising approach for tumor detection using surgical probes.

CHAPTER 1

INTRODUCTION

One goal of nuclear medicine is to detect tumors. The patient is administered a radiopharmaceutical consisting of two parts: a tumor-seeking compound which the tumor tissue either metabolizes or incorporates, and a radionuclide label. Regions where the tumor-specific compound concentrates can be identified by the gamma rays emitted by the radionuclide label, and potential tumor sites can be located using gamma-ray detectors to search for areas with high count rates. Unfortunately the tumor detection task is complicated by several factors. First, no compound is strictly tumor-specific. Instead of concentrating solely in tumor tissue, some of the compound is taken up by other tissues of the body, and some remains in the bloodstream. Second, different organs take up different concentrations of the radiopharmaceutical, leading to an inhomogeneous background level of radioactivity against which small tumors are difficult to identify. Third, because the amount of radionuclide taken up by tumors is roughly proportional to the tumor volume, small tumors are inherently more difficult to detect than large tumors. Finally, the concentration of radionuclide used must be low to minimize the patient's radiation exposure.

The conventional imaging technique for tumor detection is to use a gamma camera combined with a collimator to take a two-dimensional projection image. Figure 1-1 shows the most commonly used gamma-ray camera, the Anger camera (Anger, 1958). Due to the low concentrations of radionuclide used and the poor efficiency of the collimator, the resulting images are often photon-starved. Also, the radiation from distant sources is attenuated and scattered by intervening tissue, so both resolution and sensitivity degrade as distance from the camera increases. These effects make deep tumors smaller than 2-3 cm difficult to detect (Packer, 1984).

Figure 1-1. The Anger Camera. The collimator forms a two-dimensional image of a three-dimensional object onto the detector, a large scintillation crystal. The interaction of a gamma ray with the scintillation crystal causes a flash of light which is amplified and converted into voltage pulses by a close-packed hexagonal array of photomultiplier tubes (PMTs) behind the crystal. The position of the scintillation in the crystal is calculated by an analog circuit. Further details on the operation of Anger cameras are available in Barrett and Swindell, 1981. The arrows in the figure indicate the common practice of rotating the camera around the patient to take projection images at many different angles. These projections then can be combined to reconstruct a three-dimensional volume.

An alternative to external gamma cameras is the use of small, handheld probes. During surgery such a probe can be placed close to a suspected tumor site, nearly eliminating attenuation and scatter and improving sensitivity (Woolfenden and Barber, 1989). However, only one projection angle is available with surgical probes, and without depth discrimination variations in the distant background radiation may be mistaken for tumors near the probe. Other researchers in our group have experimented with several methods for suppressing the effect of these background variations.

Figure 1-2 shows one approach to suppressing the effects of background inhomogeneities, the dual probe. A cylindrical NaI(Tl) scintillation detector is surrounded by a second annular NaI(Tl) detector. Lead collimators define the fields-of-view of the detectors. Each scintillation crystal is coupled separately to a photomultiplier tube (PMT) by a fiberoptic lightguide. The probe is maneuvered so that the central, or inner, detector measures both the suspected tumor and the background, while the outer detector measures the background alone (Hickernell *et al.*, 1988; Hickernell, 1988; Hartsough *et al.*, 1989). As Figure 1-2 shows, the collimators of the two detectors are designed so that the detectors measure approximately the same background distribution. Therefore the background counts in the inner detector, C_{ib} , can be estimated from the counts in the outer detector, C_o , by the expression $C_{ib} \simeq FC_o$, where F is a constant that depends on the probe geometry. When the estimated background counts in the inner detector are subtracted from the total counts in the inner detector, C_i , most of the effect from the background activity is removed, leaving only the tumor signal. In simulated clinical trials using a numerical torso phantom, the dual probe showed a significant improvement in tumor detection when compared to a single-element probe (Hickernell *et al.*, 1988). The quantity $C_i - FC_o$ was the test statistic used for the dual probe, while C_i was the test statistic for the single-element probe. The background subtraction method assumes, however, that only the inner detector

Figure 1-2. Cutaway View of the Dual Probe. The central, cylindrical NaI(Tl) scintillation detector is 0.6 cm in diameter. The second, concentric NaI(Tl) scintillation detector has an outer diameter of 1.8 cm. The detectors are both 0.6 cm long and are separated by a gold annulus 0.1 cm thick. The detectors are separately coupled to photomultiplier tubes by fiberoptic light guides. Figure reproduced from Hickernell *et al.*, 1988.

receives counts from the tumor. If the outer detector also receives counts from the tumor, those tumor counts will be incorrectly interpreted as background counts and weighted and subtracted from the inner detector counts, thereby reducing the tumor signal. Also extended complex near-field structures such as blood vessels can give tumor-like responses, causing false positives.

A more recent approach is the use of imaging probes. Figure 1-3 shows a 21-element imaging probe built by our group. The detectors are 0.2cmx0.2cmx0.2cm cubes of cadmium telluride (CdTe) arranged in a 5x5 square grid with the corner detectors missing. The probe is equipped with a parallel-bore collimator made of gold, one bore per detector. The dimensions given in the figure caption were selected for use with the radioisotope ^{99m}Tc . In practice a 3Space Isotrack™ localizer (Polhemus Division, McDonnell Douglas) is glued to the back of the probe to measure the position of the probe in three-dimensional space with a resolution of about 0.16 cm. The spatial coordinates of the probe are used to identify corresponding pixels in overlapping images, enabling individual images to be combined to produce an extended-field map. The goal of using imaging probes is that the images will allow a surgeon to distinguish increases in photon count rate caused by tumors in the field of view or by normal structures, such as blood vessels, from those caused by variations in the background radiation level. The 21-element probe has been used to perform several intraoperative tumor searches.

Scope of this Dissertation

This dissertation explores a third method of background suppression: collimatorless coincidence imaging, a decidedly unconventional imaging system that images without a collimator or aperture of any kind. The heart of this dissertation is a demonstration that coincidence detection alone, without collimators, time-of-flight information, or

Figure 1-3. 21-Element Imaging Probe. A representative drawing of a prototype imaging probe which has 21 CdTe detectors, each 0.2x0.2x0.2cm, arranged in a 5x5 grid with the corner detectors missing. The center-to-center spacing between detectors is 0.30 cm. The parallel-bore collimator is made of gold and has one bore per detector. Each bore has a diameter of 0.24 cm and a length of 0.75 cm. Figure reproduced from Woolfenden and Barber, 1989.

other means of restricting the field of view, is sufficient to achieve useful imaging of objects near the detectors. The next chapter describes coincidence imaging, summarizes previous work in this area, and derives the imaging equation for collimatorless coincidence imaging. Chapter 3 describes the computer simulations performed to generate data sets similar to those that a collimatorless coincidence probe might acquire. The method by which these data were reconstructed into object representations is presented in Chapter 4. Chapter 5 is a brief description of semiconductor detector physics. Chapter 6 presents the laboratory measurements performed to verify that the results given in Chapter 4 are realistic predictions of probe performance. Chapter 7 summarizes our conclusions and suggests areas for further research.

CHAPTER 2

COINCIDENCE IMAGING

2.1 Definition of coincidence imaging

Several radionuclides used in nuclear medicine emit two photons per decay almost simultaneously, either two gamma rays in cascade or a gamma ray and a K x ray. These isotopes usually do not produce strong angular correlation between the coincident photons; unlike photon pairs produced by positron annihilation which travel in opposite directions, coincident photons may travel at any angle with respect to each other. Coincidence imaging is the attempt to measure the instances where both photons from a single decay reach a set of detectors. This constrains the possible locations for that decay to the common field-of-view of the detectors (barring scattering), as illustrated in Figure 2-1.

Isotopes suitable for coincidence imaging must fulfill several requirements. First, the energies of the two photons must be high enough that the photons can travel through tissue to the detector(s), but not so high that they penetrate the detector(s) without interacting. In practice, this restricts the usable energies to between 50 keV and 511 keV (Sharp *et al.*, 1985). Second, both photons must have a high probability of emission. Since both photons are required for a coincidence event, a low probability of emission for either photon dramatically reduces the efficiency of the system. Third, the isotope should have a half-life greater than several hours but less than a week. Shorter lifetimes require an on-site reactor or accelerator and complicate handling. Isotopes with longer lifetimes that are not rapidly excreted continue to irradiate the patient long after the imaging session, causing radiation dose concerns. Finally, the energies of the two photons should be different so that they can be distinguished by the detectors. For example, the energy difference between 170 keV and 240 keV is readily resolved by most detectors. If the photon energies can be

Figure 2-1. Common Field-Of-View of Two Collimated Detectors. If two collimated detectors receive photons from the same radioactive decay, the location of the decay must lie within that portion of their fields-of-view which overlap, indicated by crosshatching.

distinguished, then half of all accidental coincidences, where photons from different decays strike the detectors simultaneously, can be eliminated from the data by rejecting coincidence events where both photons have the same energy. On the other hand, if the two photons have very different energies, say 15 keV and 300 keV, the detector system may not be able to distinguish the higher energy from the composite event where both photons strike the same detector (315 keV in this example).

Figures 2-2 and 2-3 show the level diagrams of two suitable isotopes for coincidence imaging, ^{111}In and ^{123}I . The radioisotope ^{111}In emits two gamma rays in cascade, one at 171 keV with probability of emission of 90.2%, followed by one at 245 keV with probability of emission of 94.0%. Since the emissions are independent events, the probability of both photons being emitted is the product of their individual probabilities or 84.8%. The half-life of ^{111}In is 2.807 days (Browne and Firestone, 1986).

The radioisotope ^{123}I emits a 159 keV gamma ray with probability of emission 83.3% and a 28 keV K x ray with probability of emission 87.0%, yielding a probability of coincidence of 72.5%. Its half-life is 13.2 hours (Browne and Firestone, 1986). Since ^{111}In has a higher probability of coincidence we chose it for both simulations and laboratory work.

2.2 Earlier work in coincidence imaging

Three research groups have contributed the majority of papers on coincidence imaging using gamma rays. The first group was founded by Inge Feuerhake (later Inge Schmitz-Feuerhake). Her goal was to find a method of accomplishing three-dimensional scintigraphy in the late 1960s-early 1970s. At the time no computer algorithms existed for combining two-dimensional projection images into a three-dimensional reconstruction and no detectors were fast enough to measure the difference in the time of flight of gamma rays.

Figure 2-2. Level Diagram of ^{111}In . ^{111}In decays by electron capture (EC) to ^{111}Cd when an inner, K-shell electron combines with a proton, creating a neutron and decreasing the atomic number of the nucleus by one. There is a 90.2% probability that this event will be followed by the emission of a 171 keV gamma ray from the nucleus. The half-life of this radioactive decay is 0.12 nsec. There is a 94.0% probability that the 245 keV ^{111}Cd atom will decay to the ground state, emitting a 245 keV gamma ray. The half-life of this radioactive decay is 85 nsec. Reference: Browne and Firestone, 1986.

Figure 2-3. Level Diagram of ^{123}I . ^{123}I decays by electron capture (EC) to ^{123}Te when an inner, K-shell electron combines with a proton. There is a 83.3% probability that the nucleus will emit a 159 keV gamma ray. The half-life of this radioactive decay is 0.20 nsec. There is a 87.0% probability that an L-shell electron will drop down to fill the vacancy left in the K-shell, emitting a 28 keV K x ray in the process. The K x ray is not shown on the level diagram because it is not a nuclear transition. Reference: Browne and Firestone, 1986.

Instead Dr. Feuerhake proposed using coincidence imaging combined with collimated detectors to restrict the volume of an object that was imaged. Her earliest paper published in English (Schmitz-Feuerhake, 1970) describes phantom studies where coincidence between the 130 keV and 270 keV gamma rays of ^{75}Se was measured by the experimental setup shown in Figure 2-4. A coincidence event was recorded whenever one gamma ray was detected by a scanner moving across the top of the phantom while at the same time a second gamma ray was detected by plastic scintillators on either side of the phantom. The long focusing slit collimators on the plastic scintillators restricted the volume being imaged to a horizontal slab within the phantom, while the position of the scanner further restricted the possible decay sites to a disk. The results were unsatisfactory due to the low efficiency of the system.

In 1979 another German group updated her experiment, replacing the scanner with a scintillation camera and the plastic scintillators with NaI(Tl) detectors (von Boetticher *et al.*, 1979). Schmitz-Feuerhake and von Boetticher combined on several later papers that discussed methods of improving the efficiency of this gamma-gamma coincidence system, comparing pinhole versus parallel-hole collimators in front of the scintillation camera, optimizing the slit collimator parameters, and adding energy windowing to reduce the contribution from scattering (von Boetticher *et al.*, 1982). They also built and tested a 3D scanning configuration of seven scintillation detectors (six surrounding a central detector) all with focusing collimators focused on the same point (Helmers *et al.*, 1979, Helmers *et al.*, 1982).

A second major contributor to the literature of gamma-gamma coincidence imaging is Hiram Hart's group. In 1964 Hart proposed setting the pulse height window of a standard focusing collimator scanner to the sum of the energies of the coincidence photons, so that only coincidences would be recorded (Hart, 1965; Hart, 1968). In a subsequent

Figure 2-4. Schmitz-Feuerhake Experiment. Coincidences are measured between the scanner and the plastic scintillators. The scanner is a NaI scintillation crystal with a diameter of 12.7 cm combined with an RCA 8055 photomultiplier tube. The two plastic scintillators are 30cmx15cmx10cm. The focusing slit collimators (or magnifying collimators) consist of 11 plates of lead 0.2 cm thick pointed at a 2 cm thick section of the phantom. The phantom consists of four spheres in water. Each sphere is filled with 10 μCi of ^{76}Se . The scanning time is 10 minutes. From Schmitz-Feuerhake, 1970.

paper (Hart and Rudin, 1977) he proposed using coincidence imaging to reduce the longitudinal depth of focus of focusing collimators. Figure 2-5 shows the scanning geometry he used. Two scintillation detectors with focusing collimators were aligned at right angles. Each of the two probes was gated to respond to only one of the two ranges of gamma-ray energies of ^{75}Se : 97 to 136 keV and 265 to 280 keV. If gamma rays in these energy ranges were detected by the two detectors within 55 nsec, a coincidence event was recorded. Although the system was able to detect a cold spot of 0.03 cm^3 the scan time was prohibitively long, requiring two hours to scan a single $1.2\text{cm} \times 1.8\text{mm}$ plane using a 0.2 cm scan spacing. In later papers Hart increased the number of detectors to five arranged in an arc, doubled the channel dimensions of the collimators, and allowed all detectors to accept gamma rays of both energies (Chung *et al.*, 1980). Even with these improvements the scan time was 30 sec per pixel so that a scan of a $4.5\text{cm} \times 4.5\text{cm} \times 2.7 \text{ cm}$ volume on a 0.3 cm grid (2025 pixels) took 16.9 hours. The group has also explored triple coincidence imaging using isotopes that emit three gamma rays in cascade (Liang *et al.*, 1987; Liang, 1988).

The third group began measuring coincidences with two Anger cameras at right angles, one equipped with a parallel bore collimator and the other with a slit collimator (Powell *et al.*, 1970; Monahan *et al.*, 1972), illustrated in Figure 2-6. The detection of a gamma ray by the camera with the slit collimator determined the plane of the decay, while the detection of a second gamma ray by the camera with the multihole collimator determined the line of propagation. The intersection of the line with the plane gave an (x,y,z) coordinate for the decay. As with the previous two methods described, the coincidence counting efficiency was quite low. The total coincidence rate was about 1 count per second, and it took 60 hours to accumulate 200,000 coincidence counts. With 200,000 counts the system could detect cold defects down to 12 cm^3 in volume and specific activity ratios of at least 4:1 in a volume of 4 cm^3 .

Figure 2-5. Hart Experiment. Each detector is a 12.7 cm diameter NaI(Tl) scintillation crystal with an RCA 4525 bi-alkali photomultiplier tube. The 1500-channel lead collimator is 12.7 cm in diameter and 6.35 cm thick. The axes of the detectors are at right angles to each other, and the focal points of the two collimators are aligned to within 0.2 mm in the x direction. The phantom consists of a volume 3.2 cm x 3.8 cm x 1.9 cm containing 160 μCi of ^{75}Se with a single spherical cold defect of radius 0.3 cm. Scans performed on a 0.2 cm x 0.2 cm x 0.2 cm grid successfully indicated the presence of the cold defect, but the scanning time was 2 hours. From Hart and Rudin, 1977, © 1977 IEEE.

Figure 2-6. Monahan Experiment. Two Anger cameras are arranged at right angles. One camera is equipped with a conventional parallel bore collimator while the other has a slit collimator consisting of thirteen 1.2 cm wide slots separated by 1.2cmx7.5cm lead septa. The slit collimator determines the z coordinate of the decay, while the multihole collimator determines the (x,y) coordinate. The phantom consists of a box filled with 2600 cm³ of water containing 50 μ Ci of ⁷⁵Se and three cylinders of Lucite with diameters of 2.5 cm, 3.8 cm and 5.0 cm. The length of each cylinder is equal to its diameter. The box is located 5 cm above the slit collimator and 12 cm from the multibore collimator. Pulse-height analyzers accept events in the range from 110 keV to 300 keV. The resolving time of the coincidence circuit is 30 nsec. From Monahan *et al.*, 1972.

In 1989 Powell wrote a theoretical paper suggesting applying time-of-flight (TOF) information to coincidence imaging using the geometry shown in Figure 2-7. A coincidence event occurs when one gamma ray is detected by a collimated detector at F_0 and a second gamma ray from the same decay is detected by a small uncollimated detector at F_n . If a sufficiently short coincidence resolving time can be achieved then the difference in the TOF of the two gamma rays can be measured, and from the TOF the difference in path lengths of the two gamma rays can be obtained. The path length difference and the location of the foci, F_0 and F_n , establish the surface of a hyperboloid H_n . Ideally, the intersection of the surface of this hyperboloid with the line determined by detection of the gamma ray by the collimated detector gives the origin of the decay. Actually, the field of view of the collimated detector determines a cone of possible decay sites rather than a line, and the full width at half maximum (FWHM) of the resolving time of the TOF circuit leads to a family of possible hyperbolic surfaces. At a depth of 10 cm and the currently achievable FWHM of 500 psec, the uncertainty in the location of the event would be a volume approximated by a cylinder 1 cm in diameter and 7.5 cm in length. A good summary of all these approaches to coincidence imaging is presented by Liang, 1990.

All the methods of coincidence imaging described so far suffer from low efficiencies caused by using collimated detectors to measure one or both of the coincident photons. The heart of this dissertation is a demonstration that coincidence detection alone, without collimators, time-of-flight information, or other means of restricting the field of view, is sufficient to achieve useful imaging of objects near the detectors. The detected coincidence data contain a great deal of information about the object distribution, and this information can be recovered with a simple matrix pseudoinverse. The absence of a collimator means no trade-off between sensitivity and resolution is necessary. In the next section we derive a theoretical form for the response of a collimatorless coincidence imaging

Figure 2-7. Powell Time-of-Flight Geometry. The difference in the time-of-flight is measured between photons arriving at a detector with a multibore collimator and photons interacting with small uncollimated detectors distributed around the patient, F_n ($n=1,2,\dots,9$). The H_n are sections of hyperbolas determined by the arrival of a photon at F_0 and a photon at F_n . The depth of the source point illustrated is approximately 10 cm. From Powell, 1989.

system.

2.3 Derivation of H matrix expression

Without a collimator, the probability of a particular coincident event is determined simply by solid angles. In the following discussion "pixels" and "detectors" are synonymous, and vectors are indicated by bold typeface. To clarify the variables used in the derivation consider a two-dimensional detector array in the x-y plane viewing a three-dimensional distribution of activity (Figure 2-8). The probability that an isotropically emitted photon from location $\mathbf{R}$ reaches pixel i is directly related to the solid angle $\Delta\Omega$ subtended by pixel i :

$$P(1 \text{ phtn}) = \frac{1}{4\pi} \Delta\Omega = \frac{\epsilon^2}{4\pi} \frac{\cos^3\theta_i(\mathbf{R})}{z^2}, \quad (2.1)$$

where ϵ^2 is the pixel area, $\theta_i(\mathbf{R})$ is the angle between the normal to the i^{th} pixel and the line connecting the i^{th} pixel to location $\mathbf{R}$, and z is the vertical distance from $\mathbf{R}$ to the detector plane. In the right side of Equation 2.1 we have made the common assumption that the detector area is much smaller than the distance from the source point to the detector. We assume that the detector efficiency is unity. In coincidence imaging we require that two photons from $\mathbf{R}$ reach pixels. If we assume there is no correlation between the two photons, then the probability of coincident detection in pixels i and j is simply the product of the individual probabilities:

$$P(2 \text{ phtns}) = \frac{\epsilon^4}{16\pi^2} \frac{\cos^3\theta_i(\mathbf{R})}{z^2} \frac{\cos^3\theta_j(\mathbf{R})}{z^2}. \quad (2.2)$$

The sensitivity of coincidence imaging to source depth can be seen from the $1/z^4$

Figure 2-8. Two-Dimensional Detector Array Viewing a Three-Dimensional Distribution of Activity. The plane of the detector array is the x-y plane, and the vertical distance of a radioactive decay from the detector plane is the z coordinate. If the radioisotope used produces one photon per decay, then a decay at location **R** produces a photon which is emitted at angle θ_i . If the radioisotope emits two photons per decay, a decay at location **R** produces two photons: a photon emitted at angle θ_i and another emitted at angle θ_j . The pixels where these photons strike the detector array are labelled **i** and **j** respectively.

dependence of this expression; thus, the effects of distant background sources should be highly suppressed.

The digital imaging process can be represented by the continuous-discrete model described in Barrett *et al* (1991):

$$g_m = \int_{\infty} dx dy dz f(x, y, z) h_m(x, y, z) + n_m , \quad (2.3)$$

where g_m is the m^{th} measurement, $f(x, y, z)$ is the continuous three-dimensional object being imaged, $h_m(x, y, z)$ is the average contribution of the object point at location (x, y, z) to the m^{th} measurement, and n_m is random noise in the data. If we neglect attenuation and scatter and assume a detector efficiency of unity, h_m for our collimatorless coincidence probe is given by the right side of Equation (2.2), where m lexicographically indexes all possible combinations of detectors i and j .

Although in the real world the object volume is continuous and, as described in the next chapter, we used a continuous object volume to generate simulated data, a discrete object representation is necessary before image reconstruction from the data can occur. The continuous object $f(x, y, z)$ can be expanded in a discrete set of basis functions:

$$f(x, y, z) \simeq \sum_{n=1}^N f_n \Phi_n(x, y, z) . \quad (2.4)$$

Equation 2.4 is an approximation of the continuous object because the number of f_n in the summation is finite. As basis functions we chose a set of voxels where $\Phi_n(x, y, z)$ was unity within the n^{th} voxel and zero elsewhere. We divided the object space into a 10x10x10 grid

of voxels where the length of each side of a voxel equalled the width of one detector.

With the discrete object representation given by Equation 2.4 we can define a matrix $\mathbf{H}$ which is related to h_k in Equation (2.3) by:

$$H_{mn} = \int_{\infty} dx dy dz h_m(x, y, z) \Phi_n(x, y, z) , \quad (2.5)$$

and the continuous-discrete model given in Equation 2.3 becomes a discrete-discrete imaging model given by the matrix equation:

$$g_m = \sum_{n=1}^N H_{mn} f_n + n_m , \quad m=1,2,\dots,M , \quad (2.6)$$

where the object is represented by an $N \times 1$ column vector $\mathbf{f}$ with components f_n , the data set is represented by an $M \times 1$ column vector $\mathbf{g}$ with components g_m , and the imaging system is represented by an $M \times N$ matrix $\mathbf{H}$ with elements H_{mn} . The matrix element n_m represents random noise. If we neglect attenuation and scatter and assume a detector efficiency of unity, we can write an element of $\mathbf{H}$ as:

$$H_{mn} = H_{(ij)n} = \frac{\epsilon^4}{16\pi^2} \frac{\cos^3\theta_i(x_n, y_n, z_n) \cos^3\theta_j(x_n, y_n, z_n)}{z_n^4} , \quad (2.7)$$

where (x_n, y_n, z_n) is the n^{th} location in object space, and the other quantities are defined in the text for Equation 2.1. The analytic expression in Equation 2.7 is the model for the behavior of a collimatorless coincidence imaging system that we shall use in Chapter 4.

2.4 True versus Accidental Coincidences

Two types of coincidences are measured by any coincidence system: true coincidences and accidental coincidences. It is important to understand the difference between these types and to be able to estimate the proportion of each in a data set.

True Coincidences

A true coincidence occurs when the decay of a single atom within the ^{111}In source releases two photons in directions that cause both photons to hit detectors. This includes the possibility that both photons hit the same detector. The probability of this event is governed by Equation 2.2 and varies as the inverse distance between source and detector to the fourth power. Since a true coincidence originates from a single decay of ^{111}In , the number of true coincidences varies linearly with the source activity.

To calculate the number of true coincidences to expect in some time interval, Δt , we need to know the number of 171 keV and 245 keV photons a source will produce. Referring to the level diagram for ^{111}In , Figure 2-2, we see that there is a 90.2% probability that the 171 keV photon will be emitted, and a 94.0% probability that the 245 keV photon will be emitted. A true coincidence requires that both photons be emitted. Since the photons are emitted independently, the probability of both being emitted is the product of their individual probabilities. Therefore the number of pairs of 171 and 245 keV photons emitted in measurement time Δt is:

$$N_{171+245} = (0.902) (0.940) A \Delta t = (0.848) A \Delta t , \quad (2.8)$$

where A is the activity of the source in Becquerels. The Becquerel is the SI unit of activity, abbreviated by the letters Bq. One Becquerel equals one disintegration per second. The two photons are emitted nearly isotropically into 4π steradians. Equation 2.2 gives the fraction

of photons emitted at source point $\mathbf{R}$ that fall within the solid angles subtended by the two detectors. For a perfect coincidence detection system, the expected number of true coincidences from a point source at $\mathbf{R}$, $N_t(\text{ps})$, in time interval Δt would be:

$$N_t(\text{ps}) = N_{171+245} P(2 \text{ phtns}) = (0.848) A \Delta t \frac{\epsilon^4}{16\pi^2} \sum_{i=1}^D \sum_{j=i}^D \frac{\cos^3\theta_i(\mathbf{R})}{z^2} \frac{\cos^3\theta_j(\mathbf{R})}{z^2}, \quad (2.9)$$

where θ , $\mathbf{R}$, ϵ and z are the same as in Equation 2.1 and the summation is performed over all pairs of D detectors. For an extended source distribution Equation 2.9 becomes a volume integral:

$$N_t(\text{es}) = (0.848) \Delta t \frac{\epsilon^4}{16\pi^2} \int d^3\mathbf{R} \sum_{i=1}^D \sum_{j=i}^D \frac{\cos^3\theta_i(\mathbf{R})}{z^2} \frac{\cos^3\theta_j(\mathbf{R})}{z^2} \mathcal{A}(\mathbf{R}), \quad (2.10)$$

where $\mathcal{A}(\mathbf{R})$ is the activity per unit volume.

Accidental Coincidences

It is more likely that just one of the two emitted photons from a decay will hit a detector. Equation 2.1 shows that the probability of one photon striking a detector varies as the inverse second power of distance. An accidental coincidence occurs when a photon from one decay reaches the detector array at the same time as a photon from a different decay. Because an accidental coincidence requires that two independent decays of ^{111}In occur, the number of accidental coincidences varies as the square of the source activity. Note that the second decay can occur at the same source point as the first decay or at a different source point within the volume. Also note that the definition of arriving "at the same

time" depends on the details of the coincidence circuit. The circuit accepts events occurring within some time interval, δt , as being coincident. The longer the coincidence detection interval, the more accidental coincidences will be included in the data.

To calculate the number of accidental coincidences, we again turn to the level diagram, Figure 2-2. The number of 171 keV photons emitted in time interval δt is:

$$N_{171} = 0.902 A \delta t , \quad (2.11)$$

where, again, A is the activity of the source in Becquerels. Similarly, the number of 245 keV photons emitted in time interval δt is:

$$N_{245} = 0.940 A \delta t . \quad (2.12)$$

Accidental coincidences arise when photons from separate decays reach the detectors within the coincidence window set by the electronics. Therefore the number of accidental coincidences to expect from a point source at location $\mathbf{R}$ is the number of photons of each energy multiplied by the probability that each single photon reaches a detector, $P(1 \text{ phtn})$ from Equation 2-1:

$$N_a(\text{ps}) = N_{171} N_{245} P(1 \text{ phtn})^2 = (0.848) A^2 \Delta t \delta t \frac{\epsilon^4}{16\pi^2} \sum_{i=1}^D \sum_{j=i}^D \frac{\cos^3\theta_i(\mathbf{R}) \cos^3\theta_j(\mathbf{R})}{z^4} , \quad (2.13)$$

where Δt is the total counting time and δt is the length of the coincidence interval. Equation 2.13 assumes that accidental coincidences involving two photons with the same energy

can be identified and discarded from the data. As expected, the longer the coincidence interval, δt , the greater the number of accidental coincidences. For an extended source Equation 2.13 becomes two volume integrals:

$$N_a(es) = (0.848) \Delta t \delta t \frac{\epsilon^4}{16\pi^2} \int d^3R_1 \int d^3R_2 \sum_{i=1}^D \sum_{j=i}^D \frac{\cos^3\theta_i(R_1)}{z_1^2} \mathcal{A}(R_1) \frac{\cos^3\theta_j(R_2)}{z_2^2} \mathcal{A}(R_2) . \quad (2.14)$$

Subscripts have been added to R , the distance of the source point from the detector, z , the vertical distance to the detector plane, and $\mathcal{A}(R)$, the activity per unit volume of the source at R , to indicate that the photons causing an accidental coincidence can originate at different locations.

Next we examine the ratio of the number of true coincidences to the number of accidental coincidences. For a point source we can directly take the ratio of Equation 2.9 to Equation 2.13:

$$\frac{N_t}{N_a} = \frac{1}{A \delta t} . \quad (2.15)$$

This says that the ratio of true coincidences to accidental coincidences for a point source is independent of the detector configuration and depends only upon the activity of the source and the length of the coincidence interval. Equation 2.15 indicates that the low concentrations of radioisotope common in nuclear medicine are an asset for coincidence imaging, since a low activity, A , increases the percentage of true coincidences in the data. Similarly Equation 2.15 suggests that a coincidence circuit should be designed to use the shortest possible coincidence interval, δt , further increasing the proportion of true coincidences in the

data. The ratio of true to accidental coincidences for an extended source (Equation 2.10 divided by Equation 2.14) is more complex and is dependent on both the detector configuration and the distribution of activity in the source.

CHAPTER 3

MONTE CARLO SIMULATIONS

In order to study the behavior of a collimatorless coincidence detector array, we simulated its behavior using Monte Carlo methods. The goal of the simulation was to produce coincidence data sets similar to those that a real detector array might record. In Monte Carlo simulations, this is done by individually tracing many photons through the system from emission to detection. Whenever the probability law governing a photon's travel changes, a random intermediary output is constructed according to the new probability law and this intermediary output is used as the input to the next stage of the photon's travel. The process is repeated until the photon reaches a final output state. This technique is particularly useful for problems which are too complex for an analytic approach.

3.1 Description of probe geometry

The radioisotope modelled was ^{111}In . Figure 2-8 shows the geometry used in the simulations. The detector array is a 10x10 grid. Each detector in the array is 0.2cmx0.2cm. This geometry was chosen because we plan to build a surgical probe of these dimensions using semiconductor detectors with a row-by-column readout scheme. The probe will have a detachable collimator so that both conventional imaging and collimatorless coincidence imaging can be accomplished.

In the earliest simulations, the object volume consisted of a cube 3 cm on a side centered under the detector. In simulations of extended objects a uniform background activity was assigned to the cube, and tumors, represented by spheres, were assigned an activity ten times greater than the background. This is a reasonable contrast between tumor and background for ^{111}In -labelled monoclonal antibodies. For trials involving a point

source, the background activity was set to zero, meaning that all photons came from the same location.

3.2 Step-by-step procedure used in Monte Carlo simulations

We now describe the procedure used in generating simulated data. A point (x, y, z) in the continuous object volume was randomly chosen using uniform deviates generated by the RAN1 routine in *Numerical Recipes* (Press *et al.*, 1986). The activity at that point was compared to another random number to determine whether gamma-ray emission occurred. (These first two steps were skipped for point source simulations.) If the point emitted, a uniformly-distributed random direction was chosen for the first photon. The direction of the second photon was determined according to the expression for the angular correlation between the two cascade gamma rays of ^{111}In (Hinman *et al.*, 1964):

$$W(\theta) = 1 + A_2 P_2(\cos\theta) = -0.135 \cos 2\theta + 0.955, \quad (3.1)$$

where $W(\theta)$ is the relative probability that the second gamma ray in the cascade is emitted at an angle θ from the emission direction of the first gamma ray; A_2 is a measured coefficient in the range -0.180 ± 0.002 (we chose -0.180); and P_2 is the Legendre polynomial of degree 2. Figure 3-1 shows the difference between the correlation between photons expressed by Equation 3.1 and uncorrelated photons. For uncorrelated photons every possible direction for the second photon is equally probable, represented in Figure 3-1 by a circle. However, the orientation of the ellipse in Figure 3-1 shows that for ^{111}In the most probable direction for the second photon is perpendicular to the first.

Since we were simulating the performance of a surgical probe, we assumed that the distance between the emitting point and the probe was filled with attenuating tissue.

Figure 3-1. Correlation Between ^{111}In Photons. Suppose that the first photon in a ^{111}In decay is emitted parallel to the positive x axis. If the direction of the second photon were completely uncorrelated with the direction of the first photon, then all directions in the 4π steradians surrounding the decay site would be equally probable. Drawing contours of isoprobability for the direction of the second photon would produce nested spheres. The dotted circle in the figure represents a meridian curve through one such sphere (a cross-section through the rotation axis). However, because of the correlation between the photons of ^{111}In expressed in Equation 3.1, not all directions for the second photon are equally probable. Therefore, contours of isoprobability form nested oblate spheroids instead of spheres. The solid curve in the figure shows the cross-section through an isoprobability contour of equal magnitude to that shown for the uncorrelated case. Note that the most probable direction for the second photon is perpendicular to the first photon.

Therefore, if a photon's direction was such that it would hit the detector array, the radial distance to the detector was calculated and another random number was used to determine whether the photon was absorbed before reaching the detector. Finally, the efficiency of the detectors was set to one, so that any photon that reached a detector was assumed to be detected. Therefore, if both photons hit the detector plane, the bin (x_1, y_1, x_2, y_2) corresponding to the pair of detectors where the photons hit was incremented. Since there were 100 detectors in the array, the data vector had $(100)^2$ or 10,000 elements, representing all possible combinations of two detectors.

Figure 3-2 shows a data set generated in this manner for a point source. The data are arranged so that the large-scale 10x10 grid represents the detector where the first photon struck, (x_1, y_1) . Within each larger grid is a smaller 10x10 grid representing the detector where the second photon in the emission pair struck, (x_2, y_2) . The point source was located under detector $(x, y) = (9, 4)$ where pixel (1, 1) is the upper left corner of the detector array. The source was 0.5 cm below the detector plane.

3.3 Test of the random number generator

The Monte Carlo approach correctly incorporates Poisson noise into the data. However, it places great demands on the random number generator. A coincidence event requires ten successive calls to the generator: three to determine the source point, one to determine whether that point emits, four to determine the directions of the emitted photons, and two to determine whether the photons are absorbed before reaching the detectors. In the nomenclature used in *Numerical Recipes*, we use $k=10$ random numbers to plot points in 10-dimensional space. Most system-supplied random number routines use linear congruential generators, where a sequence of integers between 0 and $m-1$ is generated by the recurrence relation:

Figure 3-2. Two-Dimensional Display of Four-Dimensional Data Set. The figure shows a typical data set generated by simulating the response of the collimatorless coincidence probe to a point source located 0.5 cm below the detector plane. If the upper left detector in the detector array is labelled $(x, y) = (1, 1)$, then the point source is centered under detector $(x, y) = (9, 4)$. The large-scale 10x10 grid represents the detector where the first photon struck, (x_1, y_1) . Within each larger grid is a smaller 10x10 grid representing the detector where the second photon in the emission pair struck, (x_2, y_2) .

$$I_{j+1} = aI_j + c \pmod{m}, \quad (3.2)$$

and a and c are positive integers. The advantage of the linear congruential method is that it is very fast. However, the great disadvantage is that it is not free of sequential correlation on successive calls. For our example, the points generated are not uniformly distributed throughout the 10-dimensional space, but lie along 9-dimensional surfaces and there are at most $m^{1/k}$ such surfaces. This means that some physically possible combinations will never be generated. The resulting gaps in the data can lead to very puzzling results. This effect became pronounced in simulations of a collimated system described later in greater detail. Adding a collimator moves the object volume farther away from the detector plane, magnifying the effect of nonrandomness in the angle calculations. In the initial simulations, the gap in the data caused by sequential correlations in the random number generator apparently included the narrow range of angles accepted by the collimator, so no photons reached the detector plane. The RAN1 routine in *Numerical Recipes* uses three linear congruential generators to overcome sequential correlation, but it only guarantees its results for $k=2$ through 6, while our simulation routine operates at $k=10$. Therefore, we modified the Monte Carlo routine to maintain two copies of RAN1 with different seeds and selected numbers alternately from the two versions. Using two independent copies of RAN1 in effect doubled the correlation-free space to $k=12$ and appeared to correct the problem. A good description of random number generators and a series of tests for randomness are given in Morin *et al.*, 1988.

3.4 Accidental coincidences

As described in Section 2.4, it is possible for two unrelated photons originating from different decays in the object volume to strike the detector face and be accepted into

the data set as true coincidences, i.e. photons from the same disintegration. These accidental coincidences must be included in any simulation.

According to Equation 2.14 the number of accidental coincidences scales linearly with the length of the coincidence detection interval, indicating that a coincidence circuit should be designed to have the shortest coincidence interval possible. We decided that the shortest coincidence interval we could reasonably expect to implement with electronics was 1 μsec .

To incorporate accidentals we ran our simulations twice. The first time we merely counted the number of single photons that struck the detector during the 30 second imaging time. The probability law for detecting n photons over a time interval T is the Poisson probability law:

$$P_n = \frac{e^{-a} a^n}{n!}, \quad (3.3)$$

if the photons are emitted independently and the average rate of detection is a/T (Barrett and Swindell, 1981). We wanted to determine the probability that two single photons would be detected within $T=1 \mu\text{sec}$, creating an accidental coincidence. Recall that the simulated detectors were perfect, detecting all incident photons, so the probability of detecting two photons equalled the probability of two photons arriving at the detectors. By dividing the total number of single photons reaching the detector in 30 seconds by 30 seconds we obtained the average arrival rate. We multiplied this rate by $T=1 \mu\text{sec}$ to obtain the parameter a , then plugged that into Equation 3.3 with $n=2$. The result was the quantity P_2 . For the low activities commonly used in nuclear medicine, this number was quite small, indicating that few accidental coincidences have to be included in the data and that the probability of more than two photons arriving in T is approximately zero. For the Zubal phantom

simulations P_2 was less than 10^{-6} .

We then repeated the simulation. This time whenever the directions of the emitted photons resulted in only one photon from a disintegration striking the detector, a random number between zero and one was generated. If this number was less than P_2 , the fractional probability of an accidental coincidence occurring, the single photon was considered part of an accidental coincidence. The detector location where it hit was stored and the next time an object point emitted such that one (and only one) of its photons struck a detector, the histogram bin corresponding to the pair of detectors hit by the two unrelated photons was incremented.

To summarize, the Monte Carlo routine incorporated Poisson noise, the correlation between the directions of emitted photons, and attenuation. It did not include scattering. Accidental coincidences were included as described above. The detectors were assumed to be perfect. In Chapter 5 we discuss the actual characteristics of semiconductor detectors.

3.5 Object models used -- geometry and activities

Now that we have described the procedure for generating data, we present the details of the three object models used in the simulations. The simplest object model was a point source. Since any arbitrary object can be expressed as the superposition of point sources, it is useful to examine the response of the system to a point source. The response of an imaging system measured at $\mathbf{r}$ to a point source at location $\mathbf{r}_0$ is called the point response function, $p(\mathbf{r}; \mathbf{r}_0)$.

The second object model used consisted of six spheres in a cube, illustrated in Figure 3-3. The two spheres near the detector plane represented the tumors we wished to detect. The four spheres distant from the detectors simulated an inhomogeneous background

Figure 3-3. "Spheres-in-a-Cube" Object Model. The object volume consists of a 3.0 cmx3.0 cmx3.0 cm cube centered underneath the detector plane. The two 0.8 cm diameter spheres touching the detector plane represent the tumors we wish to detect, while the four 1.0 cm diameter spheres that are 9 cm from the detector plane simulate an inhomogeneous background interfering with the tumor detection task. The activity of all six spheres is 3700 Bq/cm³, which is typical for ¹¹¹In-labelled monoclonal antibody B72.3. The rest of the object volume has a uniform activity of 370 Bq/cm³ so the tumor to background contrast is 10:1.

interfering with the tumor detection task. The activity of all six spheres was 3700 Bq/cm³, which is typical for ¹¹¹In-labelled monoclonal antibodies. The rest of the volume had a uniform activity of 370 Bq/cm³. The exposure time was 30 seconds.

Since one proposed application of collimatorless coincidence imaging is in surgical probes, a logical object model to use is a realistic representation of the human body. George Zubal at Yale University kindly gave us a digital anatomic phantom (Zubal and Harrell, 1991) to use in our simulations. Figure 3-4 is a slice from this phantom. The procedure used by the Zubal group to create the torso phantom was to take transverse computed tomography (CT) images of a human torso, transfer them to a MicroVax® and display them on a screen. A senior technician from the Yale Diagnostic Radiology Department outlined the organs on the screen, then a resident identified the organs. Table 3-1 provides a list of the organs identified. Each pixel was assigned a number corresponding to an organ type. Therefore it was possible to assign different activities to different organs to provide a more realistic model of the inhomogeneous background a surgical probe would encounter. Dr. James M. Woolfenden of the University of Arizona Medical Center Nuclear Medicine Department supplied a list of typical relative concentrations in organs and typical administered dosages for monoclonal antibody B72.3 labelled with ¹¹¹In. We converted this information into probabilities of emission for each organ. These are presented in Table 3-2. Since tumor contrast rarely exceeds 10:1, we assigned tumors the same probability of emission as the liver and spleen, 7.1 times the probability of emission for skin. In each case the tumor was 0.8 cm in diameter, centered on the array and in contact with it, and the imaging time was 30 seconds.

Figure 3-4. Slice from the Zubal Torso Phantom. The torso phantom consists of 77 horizontal slices from a single patient. Each slice has 512x512 pixels. Each pixel is 1mmx1mm, and the distance between slices is 10mm. Each pixel is assigned a number corresponding to an organ type. To minimize storage space, we run-length encoded the phantom into a series of pairs of integers where the first integer was the organ type (see Table 3-1) and the second was the number of repetitions of that organ type. The activities assigned to different organs were derived from a list of typical relative concentrations in organs and typical administered dosages for monoclonal antibody B72.3 labelled with ^{111}In supplied to us by Dr. James M. Woolfenden of the University of Arizona Medical Center Department of Radiology, Division of Nuclear Medicine. The activity assigned to skin was 159.1 Bq/cm^3 ($= 4.3 \text{ nCi/ml}$), while the activity assigned to tumors was the same as the activity of the liver and spleen, 1113.7 Bq/cm^3 . Two tumor sites in the large intestine were selected. The cross-hairs in the image indicate the site of one tumor, 16 cm below the liver. The other tumor site was 3 cm below the liver. The tumors were 0.8 cm in diameter, centered on the array and in contact with it. To speed up processing the phantom was truncated to a volume $4.0\text{cm} \times 4.0\text{cm} \times 10.0\text{cm}$, centered on the tumor and extending 10 slices above it. The detector was horizontal, pointing up toward the shoulders of the phantom.

Table 3-1
Organs Identified in Zubal Phantom

Organ Code	Organ
64	skin
65	brain
66	spinal cord
67	skull
68	spine
69	rib cage & sternum
70	pelvis
71	long bones
72	skeletal muscle
73	lung
74	heart
75	liver
76	gall bladder
77	kidney
78	bladder
79	esophagus
80	stomach
81	small intestine
82	large intestine
83	pancreas
84	adrenals
85	fat
86	blood pool
87	gas volume, bowel
88	fluid volume, bowel
89	bone marrow
90	lymph node
91	thyroid
92	trachea
93	diaphragm
94	spleen
95	urine
96	feces
97	testes
98	prostate

Table 3-2
Zubal Phantom Organs Grouped by Activity

Organ Code	Organ	Probability of Emission
64,72,74,78,79,80,93,97,98	skin, muscle, etc	0.141
65	tumor	1.000
66	spinal cord	0.014
67,68,69,70,71,89	bone	0.718
73	lungs	0.113
75,76,94	liver, gall bladder & spleen	1.000
77	kidney	0.211
81,82,88,96	intestines & contents	0.183
83	pancreas	0.225
84	adrenals	0.282
85,91,92	fat, thyroid & trachea	0.099
86	blood pool	0.169
87	gas volume, bowel	0.000
95	urine	0.014

3.6 Collimator comparisons

For all three object models, we also simultaneously generated simulated data for the same detector array equipped with a conventional parallel-hole collimator, with one bore per detector. (As mentioned earlier, the proposed probe will have a detachable collimator.) Since probes outperform external imaging only for objects near the probe, the collimator for a probe can be designed for high efficiency. The expression for the efficiency of a collimator is:

$$\frac{\Omega}{4\pi} = \frac{\alpha_{pf} D_b^2}{16 L_b^2}, \quad (3.3)$$

where α_{pf} is the packing fraction, D_b is the diameter of the collimator bore, and L_b is the length of the bore (Barrett and Swindell, 1981). The packing fraction is the fraction of the detector not covered by collimator material, i.e. the open fraction. For square detectors of side ℓ and circular collimator holes, α_{pf} is given by:

$$\alpha_{pf} = \frac{\text{Open area}}{\text{Total detector area}} = \frac{\pi D_b^2}{4 \ell^2}. \quad (3.4)$$

We chose $L_b = 2$ cm and $D_b = 0.18$ cm, and our detectors were 0.2 cm on a side, yielding an efficiency of 3.22×10^{-4} . The face of the collimator was placed where the collimatorless detector array had been, and the same photons were used for both systems. The collimator was assumed to be perfect, i.e. no septal penetration was included. The photon's path was extended down the collimator bore. If the photon did not collide with the collimator walls, the number of counts in the detector where it struck was incremented. Note that in the collimator case the data vector has only 100 elements, one for each detector, due to the absence

of the coincidence requirement.

The next chapter describes the method used to reconstruct an estimate of the original object from the coincidence data sets and shows reconstructions of data sets obtained using the three object models discussed in Section 3.5.

CHAPTER 4

RECONSTRUCTION OF OBJECT FROM DATA

In Chapter 3 we described how we used Monte Carlo simulations to generate data sets "typical" of those we might collect from a collimatorless coincidence imaging system. In this chapter we describe how the data were used to reconstruct estimates of the original object distributions. The inverse problem of trying to determine what object distribution produced a particular data set is a common task in medical imaging. When an imaging system produces the same data set for different object distributions, the problem of reconstructing the original object is labelled ill-posed. To determine the utility of collimatorless coincidence imaging we must determine how badly ill-posed the reconstruction task is. We will do this by using eigenanalysis to examine the matrix that describes the imaging system.

4.1 Sizes of $\mathbf{f}$, $\mathbf{g}$ and $\mathbf{H}$

Recall the discrete form of the imaging equation first introduced in Chapter 2:

$$g_m = \sum_{n=1}^N H_{mn} f_n + n_m, \quad m=1,2,\dots,M, \quad (2.6)$$

where f_n is a component of the $N \times 1$ column vector $\mathbf{f}$ representing the object, g_m is a component of the $M \times 1$ column vector $\mathbf{g}$ representing the data collected, and H_{mn} is an element of the $M \times N$ matrix $\mathbf{H}$ representing the imaging system. For simplicity we will discard the term for random noise, n_m , in the rest of this discussion. Chapter 2 mentioned the decision to divide the object space into a $10 \times 10 \times 10$ grid for reconstruction. Therefore the number of

components in the object vector $\mathbf{f}$ is $N=1000$. Chapter 3 described the detector configuration used in the simulations, 100 uncollimated detectors arranged in a 10×10 array. Therefore the number of components in the measured data $\mathbf{g}$ is the number of ways two photons can hit 100 detectors. Since we include the possibility that both photons from a decay hit the same detector, the number of possibilities is 100^2 or $M=10000$. Therefore the $M \times N$ matrix describing the imaging system, $\mathbf{H}$, is 10000×1000 .

4.2 Inverting the system matrix $\mathbf{H}$

Let us write Equation 2.6 in vector form:

$$\mathbf{g} = \mathbf{H} \mathbf{f} . \quad (4.1)$$

We have discarded the noise term to simplify the discussion. This form makes it clear that the imaging equation defines a set of M linear equations with N unknowns. Our goal is to recover the original object distribution, $\mathbf{f}$, from the data or if that is not possible an estimate of that object, $\hat{\mathbf{f}}$. Looking at Equation 4.1, our first instinct might be to multiply both sides of the equation by the inverse of the system matrix, $\mathbf{H}^{-1}$:

$$\mathbf{H}^{-1} \mathbf{g} = \mathbf{H}^{-1} \mathbf{H} \mathbf{f} = \mathbf{f} . \quad (4.2)$$

However, as discussed above, the $\mathbf{H}$ matrix for collimatorless coincidence imaging is not square, so it has no inverse. But we can form a square matrix by multiplying $\mathbf{H}$ by its transpose, $\mathbf{H}^t$. Multiplying both sides of Equation 4.1 by $\mathbf{H}^t$ yields:

$$\mathbf{H}^t \mathbf{g} = \mathbf{H}^t \mathbf{H} \mathbf{f} . \quad (4.3)$$

The matrix $H^t H$ is a square, Hermitian matrix with 1000×1000 elements. A Hermitian matrix is a matrix which equals its conjugate transpose. The importance of $H^t H$ being Hermitian is explained in the next section. Now we want to multiply both sides of Equation 4.3 by the inverse of $H^t H$:

$$(H^t H)^{-1} H^t g = f . \quad (4.4)$$

But $H^t H$ is invertible only if the rank of $H^t H$, R , the number of linearly independent columns, equals the total number of columns ($R=N=1000$). For collimatorless coincidence imaging $R < N$, as will be shown in the next section. Therefore we need a way to determine the rank of $H^t H$, to identify the portion of $H^t H$ that is linearly independent and discard the rest, leaving an invertible matrix. To accomplish this we must first identify the eigenvectors and eigenvalues of $H^t H$. The eigenanalysis of the square matrix $H^t H$ is known as the singular-value decomposition (SVD).

4.3 Singular-value decomposition of $H^t H$

A nonzero vector $\mathbf{u}$ in N -space is an eigenvector of the matrix $H^t H$, if $H^t H$ multiplied by $\mathbf{u}$ is a scalar multiple of $\mathbf{u}$. The scalar multiple is called an eigenvalue. There are N eigenvectors for the $H^t H$:

$$H^t H \mathbf{u}_n = \lambda_n \mathbf{u}_n , \quad n = 1, 2, \dots, N \quad (4.5)$$

where λ_n represents the n^{th} eigenvalue and $\mathbf{u}_n$ is the n^{th} eigenvector. Since $H^t H$ is 1000×1000 , it has 1000 eigenvectors, each of which has 1000 components. We used the analytic expression for H from Equation 2-4 to evaluate H on a $10 \times 10 \times 10$ grid of equally

spaced object points, where the spacing between points equalled the width of a detector, 0.2 cm. The x-y grid points were located at the centers of detectors, and the z points started at 0.1 cm from the detector plane. Then we calculated $H^t H$ and used the SVDCMP routine from *Numerical Recipes* (Press *et al.*, 1986) to calculate the λ_n and u_n .

The matrix $H^t H$ is Hermitian. Hermitian matrices play an important role in eigenanalysis. First, every eigenvalue of a Hermitian matrix is real, and the λ_n form a discrete spectrum. The spectrum of λ_n for $H^t H$ is graphed on a semi-log scale in Figure 4-1. The eigenvalues are arranged in order of decreasing magnitude. The abrupt falloff after the first 100 values indicates that the rank of $H^t H$ is $R=100$. The fact that the falloff occurs after λ_{100} probably reflects the fact that we chose an array with 100 detectors, so we cannot expect more than 100 independent contributions to the data. On the other hand, it might reflect our decision to divide the object space into 10 planes of 100 pixels each. Due to the $1/z^4$ dependence of H , the number of pixels in the plane nearest to the detector might affect the rank of $H^t H$.

A second important feature of Hermitian matrices is that the eigenvectors of a Hermitian matrix form a complete basis (Strang, 1980). Therefore any $N \times 1$ column vector can be expanded in terms of the eigenvectors of the $N \times N$ Hermitian matrix $H^t H$. Specifically, any object vector, f can be represented by:

$$f = \sum_{j=1}^N \alpha_j u_j , \quad (4.6)$$

where the $\{\alpha_j\}$ are coefficients defined by the scalar product of the j^{th} eigenvector u_j with the object vector f .

Figure 4-1. The Eigenvalue Spectrum for $H^t H$. The eigenvalues λ_n of $H^t H$ are arranged in order of decreasing magnitude and graphed on a semi-log scale. The analytic expression for system matrix H given in Equation 2.4 was evaluated on a $10 \times 10 \times 10$ grid of equally spaced object points, where the spacing between points equalled the width of a detector, 0.2 cm. The x-y grid points were located at the centers of detectors, and the z points started at 0.1 cm from the detector plane. Then we calculated $H^t H$ and used the SVDCMP routine from *Numerical Recipes* (Press *et al.*, 1986) to calculate the λ_n . Since $H^t H$ is a Hermitian matrix, the eigenvalues are all real and form a discrete spectrum. The abrupt falloff in the spectrum after λ_{100} indicates that the rank of $H^t H$ is $R=100$. The change in slope around component 600 indicates the limit of machine precision.

4.4 Discussion of the Eigenvectors of $H^t H$

Eigenvectors are also known as eigenobjects, since they represent physical objects which are reproduced exactly by $H^t H$ except for a scale factor, λ . Figure 4-2 shows the 10 eigenvectors with the largest eigenvalues, u_1 through u_{10} .

One feature observable in all eigenvectors is that the high-spatial-frequency content of the eigenvector decreases as the distance from the detector plane increases. Since eigenvectors represent objects that can be imaged unchanged, the loss of high frequencies indicates that less detail is resolvable the farther an object is from the detector plane. We also see instances of paired eigenvectors, two eigenvectors with equal eigenvalues that are identical except for a ± 90 degree rotation. See, for example, u_2 and u_3 , u_7 and u_8 , and u_9 and u_{10} . Because the eigenvalues of these pairs are exactly equal, the ordering within a pair is arbitrary.

Figures 4-3 through 4-12 show the next 100 eigenvectors. In general we see that as the eigenvector number increases, the high frequency content of the plane nearest the detectors also increases, culminating in eigenvector u_{100} , the last row of Figure 4-11. The 10×10 checkerboard pattern in the first plane of u_{100} corresponds to the Nyquist condition, that the highest frequency a system can image is the inverse of twice the minimum distance between detectors.

Eigenvector u_1 is unique in having negative values for all of its components. In the first 100 eigenvectors, 85 have values which are symmetric about zero, e.g. u_7 has values ranging from -0.2297 to 0.2297 in the first plane, from -0.0104 to 0.0104 in the second plane, etc. (as does its pair, u_8). Eigenvector u_5 is the first eigenvector whose values are not distributed symmetrically about zero. It has values ranging from 0.2421 to -0.1391 in the first plane, 0.0121 to -0.0071 in the second plane, and its bottom three planes have all negative values.

Figure 4-2. The 10 Eigenvectors with the largest Eigenvalues, u_1 through u_{10} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Each subimage has been separately normalized so that its lowest value is assigned to the lowest grey level, $GL = -128$ or black, and its peak value is assigned the highest grey level, $GL = 127$, which is white. Without this renormalization, the $1/z^4$ factor in H would cause all but the first column to be assigned to the same grey level, $GL = 0$.

Figure 4-3. Eigenvectors u_{11} through u_{20} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

Figure 4-4. Eigenvectors u_{21} through u_{30} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4.2.

Figure 4-5. Eigenvectors u_{31} through u_{40} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4.2.

Figure 4-6. Eigenvectors u_{41} through u_{50} . Each row is an eigenvector. The ten sub-images in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4.2.

Figure 4-7. Eigenvectors u_{51} through u_{60} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

Figure 4-8. Eigenvectors u_{61} through u_{70} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

Figure 4-9. Eigenvectors u_{71} through u_{80} . Each row is an eigenvector. The ten sub-images in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

Figure 4-10. Eigenvectors u_{81} through u_{90} . Each row is an eigenvector. The ten sub-images in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

Figure 4-11. Eigenvectors u_{91} through u_{100} . Each row is an eigenvector. The ten sub-images in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

Figure 4-12. Eigenvectors u_{101} through u_{110} . Each row is an eigenvector. The ten subimages in each row correspond to the ten x-y planes where we evaluated H , from $z = 0.1$ cm in the left column to $z = 1.9$ cm in the right column. Subimages normalized as in Figure 4-2.

The second hundred eigenvectors have the same number of pairs as the first hundred (25), and one more eigenvector whose values are not symmetrically distributed about zero. However, several interesting changes occur. The first, and least visible change, is that the plane with the greatest magnitude shifts from the first plane to the second plane. Although there is a general trend that, as the eigenvector number increases, the plane of greatest magnitude shifts downward, the transition is not abrupt. For instance, eigenvectors u_{326} and u_{327} have greatest magnitudes in the fourth plane, but u_{328} and u_{329} have greatest magnitudes in the third plane.

A second interesting feature is that the patterns in eigenvectors u_{101} to u_{110} are similar to patterns in the first 100 eigenvectors. Compare u_{101} , the top row of Figure 4-12, to eigenvector u_1 , the top row of Figure 4-2. Note that in the plane nearest the detectors (left column) the two eigenvectors have the same pattern. However, all deeper x-y planes of u_{101} undergo a polarity reversal. High grey levels (white) become low grey levels (black) and low grey levels become high. This is because all the magnitudes in u_1 are negative, while only the magnitudes in the first plane of u_{101} are negative. Therefore u_{101} boosts the first plane of u_1 while counteracting or subtracting its effects in lower planes. Examining the rest of the eigenvectors in Figure 4-12, we see that there is always a polarity reversal between the plane nearest the detectors and the deeper planes. However, the second plane is not always the same polarity as the second plane of the corresponding eigenvector in Figure 4-2. For instance, compare the sixth rows, eigenvectors u_{106} and u_6 . Pairs are harder to decipher. If we recall that the order within a pair is arbitrary then we might say that eigenvectors u_{108} and u_7 correspond and have the same polarity in the second plane, while eigenvectors u_{107} and u_8 have opposite polarities in the second plane. An alternative interpretation is the u_{107} and u_{108} are versions of u_7 and u_8 , respectively, rotated 90 degrees counter-clockwise. Then they have the same polarities as u_7 and u_8 in the second

plane.

4.5 Reconstruction of Object Distributions

The eigenvectors of the $M \times M$ square Hermitian matrix HH^t can be written:

$$HH^t \mathbf{v}_j = \lambda_j \mathbf{v}_j, \quad j = 1, 2, \dots, M \quad (4.7)$$

where HH^t is 10000×10000 and has 10000 eigenvectors, each of which has 10000 components. These eigenvectors form a complete basis for any $M \times 1$ vector, so that any data vector $\mathbf{g}$ can be expanded in the form:

$$\mathbf{g} = \sum_{j=1}^M \beta_j \mathbf{v}_j, \quad (4.8)$$

where $\mathbf{v}_j$ is the j^{th} eigenvector of HH^t , and the $\{\beta_j\}$ are coefficients defined by the scalar product of the j^{th} eigenvector $\mathbf{v}_j$ with the data vector $\mathbf{g}$. If we operate on Equation 4.5 from the left with H and compare the result to Equation 4.7, we can derive the relationship between the $\mathbf{v}_j$ and the $\mathbf{u}_j$:

$$\mathbf{v}_j = \lambda_j^{-1/2} H \mathbf{u}_j. \quad (4.9)$$

If we use Equation 4.9 to substitute for $\mathbf{u}_j$ in Equation 4.6 we obtain an expression for the object vector in terms of the $\mathbf{v}_j$ and the data vector $\mathbf{g}$:

$$\mathbf{f} = \sum_{j=1}^N \frac{\beta_j \mathbf{v}_j}{\lambda_j}. \quad (4.10)$$

Recall from Figure 4-1 that the λ_j for $j > 100$ are nearly zero. In the absence of noise this does not cause a problem in evaluating Equation 4.10 because the β_j are proportional to the λ_j and the quantity β_j/λ_j remains finite as λ_j approaches zero. However, when noise is present in the data the β_j include expansion coefficients for the noise, and as λ_j approaches zero the quantity β_j/λ_j "blows up", leading to large noise amplification in the reconstruction. The solution is to write the expansion of the object vector $\mathbf{f}$ given in Equation 4.6 as two sums:

$$\mathbf{f} = \sum_{j=1}^R \alpha_j \mathbf{u}_j + \sum_{j=R+1}^N \alpha_j \mathbf{u}_j. \quad (4.11)$$

The first sum represents the part of the object vector $\mathbf{f}$ that contributes to the data and is called the measurement component. The second sum corresponds to the part of $\mathbf{f}$ that does not contribute to the data, and is called the null component. In order to use eigenanalysis to reconstruct the object, the null component of the object must be discarded. However, without the null component we can reconstruct only an estimate of the object, $\hat{\mathbf{f}}$. This estimate is given by the expression:

$$\hat{\mathbf{f}} = \sum_{j=1}^R \frac{\beta_j \mathbf{v}_j}{\lambda_j}, \quad (4.12)$$

and $\hat{\mathbf{f}}$ is called the SVD pseudoinverse reconstruction of the object. An advantage of this

method of reconstruction compared to iterative methods is its speed. The v_j and λ_j can be calculated and stored beforehand. Then reconstruction consists of simple multiplications, divisions and additions and can be accomplished in a few seconds.

4.6 Point source results

Figure 4-13 shows the SVD pseudoinverse reconstruction of a point source centered underneath the detector plane. The point source is at $z = 0.05$ cm, the same depth as the plane in the reconstruction closest to the detector plane. The data set was obtained using the Monte Carlo procedure described in Chapter 3. The object distribution then was reconstructed using Equation 4.12. The figure contains two versions of the same 10 x-y planes in the reconstruction. In the top two rows, each subimage has been separately normalized so that its lowest value is assigned to the lowest grey level, $GL = -128$ or black, and its peak value is assigned the highest grey level, $GL = 127$, which is white. This display emphasizes the shape of the reconstructed point source in each plane. The bottom two rows show the same data when all 10 planes have the same normalization. This display emphasizes that the intensity in the plane nearest the detector far outweighs the intensities in all lower planes. Figure 4-14 shows the SVD reconstruction of a point source centered under the detector plane at $z = 0.50$ cm, the depth of the deepest plane in the reconstruction. As in Figure 4-13 the same reconstruction is presented with two different normalizations. By comparing Figures 4-13 and 4-14 we see that the plane nearest the detector always has the highest intensity regardless of where the point source is located. However, the lateral extent of the image of the point source in the plane nearest to the detectors gives an indication of how deep the point source is. Therefore if we know that the object is a point source we can determine its depth from the reconstruction.

Figure 4-13. SVD Reconstruction of a Point Source Centered under the Detector Plane at $z = 0.05$ cm. The data set was obtained using the Monte Carlo procedure described in the text. The SVD reconstruction of the object was calculated using Equation 4.12. This figure contains two versions of the same 10 x-y planes in the reconstruction. In the top two rows, each subimage has been separately normalized so that its lowest value is assigned to the lowest grey level, $GL = -128$ or black, and its peak value is assigned the highest grey level, $GL = 127$, which is white. The bottom two rows show the same data when all 10 planes have the same normalization. In both cases, the top left corner is the plane nearest the detector plane, $z = 0.05$ cm, and is the plane containing the point source. The subimage to the right is the next deepest plane at $z = 0.10$ cm and the sequence continues left to right. Each succeeding plane is 0.05 cm farther away so the last plane is at $z = 0.50$ cm.

Figure 4-14. SVD Reconstruction of a Point Source Centered under the Detector Plane at $z = 0.50$ cm. The data set was obtained by simulating a point source centered underneath the detector plane at a depth of $z = 0.50$ cm using the Monte Carlo procedure described in the text. The object distribution was reconstructed using Equation 4.12. This figure contains two versions of the same reconstruction. The differences between the two versions and details of how they are displayed are described in Figure 4.13. The point source is in the deepest plane, the last plane in each sequence.

In Figure 4-15 the magnitudes of the 100 pixels in each x-y plane of the two point source reconstructions have been summed. The top line of the graph is the plot of the pixel sums in the reconstruction displayed in Figure 4-13, while the bottom line of the graph is the plot for the reconstruction shown in Figure 4-14. The intervening lines show the results when the point source is located in the other eight x-y planes. Note that regardless of the location of the point source, the pixels in the plane closest to the detector have the greatest magnitudes. Therefore if we do not know what object was imaged, we cannot determine the depth of the object from the reconstruction. Since we cannot obtain depth information there is no point in displaying the x-y planes separately. In the remaining reconstructions we sum the corresponding pixels in the ten x-y planes to yield a single two-dimensional 10x10 image.

4.7 Spheres-in-a-cube results

Figure 4-16 presents the reconstruction of a data set generated using the object distribution in Figure 3-3, the "spheres-in-a-cube" model. There are two spheres touching the detector plane representing tumors we want to detect, and the four distant spheres simulating a nonuniform background complicating the detection task. The rest of the volume has a uniform activity which is one-tenth the activity of the spheres. The top image in Figure 4-16 is a surface map showing the pixel values in a geometric projection of the two tumors along a line perpendicular to the detector plane. It shows what the surface map of a perfect, noise-free reconstruction of the tumors would look like. To create the lower left image, corresponding pixels in the ten x-y planes of the SVD reconstruction were summed together into a single two-dimensional 10x10 image, then the grey-level value of each pixel was plotted. The grey levels were normalized so that the peak grey level in the projection image equalled the peak grey level in the SVD reconstruction. The lower right image is a

Figure 4-15. Sum of Pixels within each X-Y Plane in Reconstructions of Point Sources. The top curve in the graph was derived from the reconstruction of a point source at $z = 0.05$ cm which was displayed in Figure 4-13. The pixels in each of the subimages in Figure 4-13 were summed and the sum was plotted on a log scale. The bottom curve in the graph was derived from the reconstruction displayed in Figure 4-14, a point source at $z = 0.50$ cm where again the pixels in each plane of the reconstruction were summed and plotted on a log scale. The other eight curves are from reconstructions of point sources located in the intervening eight planes of the reconstruction.

Figure 4-16. Reconstruction of the Object Distribution in Figure 3-3. Two imaging systems were simulated using the Monte Carlo procedure outlined in the text: (1) coincidence imaging without a collimator and (2) non-coincidence imaging with a parallel-bore collimator. The collimator bores were 2.0 cm long and 0.18 cm in diameter, yielding an efficiency of 3.22×10^{-4} . The object model was described in Figure 3-3. The exposure time for both data sets was 30 seconds. The top image is a surface map of the pixel values in a geometric projection of the two spheres nearest the detector plane. It represents what the surface map of an ideal, noise-free reconstruction of the two "tumors" would look like. The lower left image is a surface map of the SVD reconstruction of the collimatorless coincidence data set, where the 10 x-y planes of the reconstruction are summed into one image. The surface maps are normalized so that the peak grey level in the noise-free image equals the peak grey level in the SVD reconstruction. The two tumors are clearly visible in the SVD reconstruction. The lower right image is a map of the data set obtained when non-coincidence imaging using a parallel-hole collimator is simulated. The same normalization is used as in the other two images. In the collimator surface map the tumors cannot be distinguished from the background.

surface map of the data set obtained when the parallel-hole collimator described at the end of Chapter 3 is used instead of collimatorless coincidence imaging. In the collimatorless coincidence image the $1/z^4$ falloff in the number of coincident photons greatly suppresses the contribution from the four distant sphere, leaving the two nearby tumors clearly visible. However in the collimated system the tumors are indistinguishable from the background. Even if the number of photons in the image is increased by either doubling the activity or the exposure time, the tumors remain indistinguishable.

4.8 Zubal phantom results

Figure 4-17 presents surface maps of the grey levels in reconstructions of tumors at two locations near the liver in the Zubal phantom. In both cases the tumor was 0.8 cm in diameter, centered on and touching the detector plane, and the exposure time was 30 seconds. The collimatorless coincidence array was positioned looking upward toward the liver. The liver complicates tumor detection because it absorbs a concentration of radioisotope equal to or greater than the uptake of tumors. When conventional probes are pointed toward the liver, the contribution from the liver can obscure signals from tumors. The first location we chose in the phantom was a relatively homogeneous region of the large intestine near the cecum, 16 cm below the liver. The activity in the intestine was 207 Bq/cm^3 , five and a half times less than the activity in the liver and tumor. The top left surface map in Figure 4-17 shows the grey level values in the SVD pseudoinverse reconstruction. The tumor is clearly visible. We simultaneously generated simulated data for the same detector array equipped with a parallel-hole collimator, with one bore per detector. In the corresponding surface map for a parallel-hole collimator, at top right, the tumor is not distinguishable. Examining the data we find that 35% of the coincident hits on the collimatorless coincidence detectors came from the tumor, while just 8% of the photons reaching the colli-

Collimatorless Coincidence

Parallel-Hole Collimator

Tumor 16cm Below Liver

Tumor 16cm Below Liver

Tumor 3cm Below Liver

Tumor 3cm Below Liver

Figure 4-17. Zubal Phantom Results. Two imaging systems were simulated using the Monte Carlo procedure outlined in the text: (1) coincidence imaging without a collimator and (2) non-coincidence imaging with a parallel-bore collimator. The collimator bores were 2.0 cm long and 0.18 cm in diameter, yielding an efficiency of 3.22×10^{-4} . For both systems the object distributions were the same, described in Figure 3-4. The imaging time was 30 seconds. The two images on the left side of the page are the collimatorless coincidence results, while the images on the right are the collimator results. The upper left image is a map of the grey level values in the reconstruction of the tumor which was 16 cm below the liver. The tumor is clearly visible in the center of the map. Thirty-five percent of the coincident hits on the detector came from the tumor. The upper right image is a map of the grey level values in the image formed by the collimator. The tumor cannot be distinguished from the background. Only 8% of the photons reaching the collimated detector originated in the tumor. The lower left image is a map of the grey level values in the reconstruction of the tumor which was 3 cm below the liver. Despite the proximity of the liver, which has the same activity as the tumor, the tumor is clearly visible. Twenty-eight percent of the coincident hits came from the tumor. Finally, the lower right image is a map of the grey levels in the image formed by the collimator. The tumor is not visible which is not surprising since only 1% of the photons reaching the collimated detector came from the tumor.

mated detectors originated in the tumor.

The second location in the phantom was just 3 cm below the liver. The results are shown in the bottom two surface maps in Figure 4-17. Because the tumor is so close to the liver, and the activity of the tumor matches the liver activity, the tumor's presence is masked by the liver activity in the parallel-hole collimator simulation (bottom right). However, the tumor is clearly visible in the collimatorless coincidence simulation (bottom left), illustrating the ability of collimatorless coincidence imaging to suppress contributions from background sources. This result is confirmed by the fact that 28% of the coincident hits on the collimatorless coincidence detectors came from the tumor, while only 1% of the photons reaching the collimated detectors came from the tumor.

CHAPTER 5

SEMICONDUCTOR DETECTOR PHYSICS

The simulations described in Chapter 3 assumed that the detectors were perfect. In a perfect detector: (1) all incident photons interact with the semiconductor crystal and (2) the energy of all incident photons can be correctly identified, i.e. the detector always produces an output pulse that is proportional to the energy of the incident photon. Real-world detectors are far from perfect. In this chapter we will discuss some of the physics that governs the behavior of semiconductor detectors and describe how these characteristics affect the imaging task.

5.1 Semiconductor materials

A semiconductor is defined as a substance whose valence band and conduction band are separated by a narrow region of forbidden energies, as shown in Figure 5-1. The energy difference between the top of the valence band and the bottom of the conduction band is called the gap energy. The semiconductor material we used, cadmium telluride (CdTe), has a gap energy $\mathcal{E}_g$ of 1.47 eV at room temperature (Sakai, 1982; Siffert, 1983). At zero temperature all the electrons in a semiconductor are in the valence band, so the semiconductor is an insulator. At higher temperatures some of the valence band electrons acquire enough thermal energy to be excited across the forbidden band to the conduction band. When an electron is promoted from the valence band to the conduction band, it leaves behind a localized region of positive charge in the valence band, called a "hole". The electron in the conduction band and the hole left behind are referred to as an electron-hole pair.

Figure 5-1. Band Structure of Semiconductors. Semiconductors are materials with a narrow separation between their valence and conduction bands. The figure shows a graph of energy, $\mathcal{E}$, versus wavevector, k , for a typical direct gap semiconductor such as CdTe. The energy difference between the top of the valence band and the bottom of the conduction band is called the gap energy, $\mathcal{E}_g$. For CdTe E_g is 1.47 eV (Sakai, 1982; Siffert, 1983).

A gamma ray interacts with a semiconductor crystal by transferring its energy to one of the tightly bound inner electrons of an atom in the crystal lattice. The interaction almost always takes place with an electron in the K shell because the most tightly bound electrons have the highest probability of photoelectric absorption. After absorbing the gamma ray, the electron is ejected from the atom with kinetic energy $\mathcal{E}_e$:

$$\mathcal{E}_e = \mathcal{E}_\gamma - \mathcal{E}_b , \quad (5.1)$$

where $\mathcal{E}_\gamma$ is the energy of the incident gamma ray, and $\mathcal{E}_b$ is the binding energy of the electron. The main gamma rays emitted by ^{111}In have energies of 171 keV or 245 keV, while the binding energy of inner electrons in CdTe is about 40 keV, so the ejected electron has a kinetic energy in excess of 130 keV. As the ejected electron travels through the crystal it loses energy by three competing processes: (1) creating electron-hole pairs by promoting electrons from the valence band across the energy gap to the conduction band, (2) nonradiatively transferring energy to the crystal lattice, creating phonons, and (3) radiatively transferring energy by the creation of photons in the infrared. We will concentrate on the first process since it is the key to how semiconductors operate as detectors.

As the high-energy photoelectron travels in the crystal it undergoes Coulomb collision with the lattice, imparting enough energy to an electron in the valence band to raise it into the conduction band, thus creating an electron-hole pair. Because only a fraction of the photoelectron's energy is expended in the collision, the photoelectron is still highly energetic afterwards. Therefore it can repeat the process of transferring energy to valence-band electrons through collisions with the lattice many times, creating many electron-hole pairs. In addition, the electrons the photoelectron promotes into the conduction band may have enough kinetic energy to create electron-hole pairs too.

For CdTe the number of electron-hole pairs produced by the photoelectric absorption of a gamma ray with energy $\mathcal{E}_\gamma$ is:

$$N_{e-h} = \frac{\mathcal{E}_\gamma}{4.43 \text{ eV}} , \quad (5.2)$$

(Siffert, 1983; Sakai, 1982). As mentioned before, $\mathcal{E}_\gamma$ for ^{111}In is either 171 keV or 245 keV, so the interaction of one ^{111}In photon with a CdTe semiconductor crystal creates tens of thousands of electron-hole pairs. The difference between the 4.43 eV factor in Equation 5.2 and the CdTe gap energy of 1.47 eV reflects the presence of the competing processes listed previously.

5.2 The motion of electron-hole pairs

The motion of a conduction-band electron through the crystal is a series of scatters in random directions due to disruptions in the lattice periodicity from: (1) impurities, (2) phonons, and (3) other electrons. Holes move through the crystal too. The position of a hole changes when an neighboring electron moves into the hole, neutralizing the positive charge there, but creating a new area of positive charge at the former site of the electron.

When a semiconductor is operated as a detector, electrodes are placed on opposite sides of the crystal and a constant electric field, $\mathbf{E}$ is applied across the crystal. All the electrons in the crystal experience a force $\mathbf{F}$:

$$\mathbf{F} = -e\mathbf{E} , \quad (5.3)$$

where e is the charge of the electron. In the valence band only electrons adjacent to holes

can respond to F . Those electrons also move in a direction opposite to the applied electric field, filling holes while creating holes at their former locations. The net effect is a migration of negative charges (electrons) in the $-E$ direction and positive charges (holes) in the $+E$ direction.

We can describe the motion of electrons and holes under the influence of F by Newton's Second Law, force equals mass multiplied by acceleration, if we use the effective mass, m^* :

$$F = m^*a , \quad (5.4)$$

where m^* is:

$$m^* = \hbar^2 \frac{1}{d^2\mathcal{E}/dk^2} . \quad (5.5)$$

The symbol $\hbar$ is Planck's constant divided by 2π , and $d^2\mathcal{E}/dk^2$ is the curvature of the conduction band for the electron effective mass and the curvature of the valence band for the effective mass of holes. As shown in Figure 5-1, the curvature of the valence band is often less than the curvature of the conduction band, so the effective mass of holes is larger than the effective mass of electrons.

Equating Equations 5.3 and 5.4 we see that the acceleration, a_e , experienced by electrons is:

$$a_e = -\frac{eE}{m_e^*} = -\frac{eE}{\hbar^2} \left[\frac{d^2\mathcal{E}}{dk^2} \right]_c . \quad (5.6a)$$

The subscript c on $d^2\mathcal{E}/dk^2$ indicates that the curvature refers to the curvature of the conduction band.

Similarly, the acceleration experienced by holes is:

$$a_h = -\frac{eE}{m_h^*} = -\frac{eE}{\hbar^2} \left(\frac{d^2\mathcal{E}}{dk^2} \right)_v, \quad (5.6b)$$

where the subscript v on $d^2\mathcal{E}/dk^2$ indicates that the curvature refers to the curvature of the valence band.

5.3 Drift velocity

The effect of the applied electric field on electrons is to superimpose a drift toward the positive electrode onto the random changes in direction caused by scatters. If the average time between scatters is τ_{sc} , then an electron's average velocity in the direction of $-E$ is $a_e \tau_{sc}$. This velocity is called the drift velocity, v_d . The drift velocity for electrons is:

$$v_{d(e)} = a_e \tau_{sc} = \frac{e E \tau_{sc}}{m_e^*}, \quad (5.7)$$

substituting for a_e using Equation 5.6a. Since the lattice structure and fraction of impurities govern the number of scatters the electron experiences, τ_{sc} differs for different semiconductor materials, and each semiconductor has a characteristic drift velocity. This characteristic velocity is often described in terms of the mobility of the charge carriers, μ :

$$v_d = \mu E = \mu \frac{V_0}{L}, \quad (5.8)$$

where V_0 is the voltage difference across the detector electrodes, and L is the length of the detector. For CdTe the mobility of electrons in the conduction band, μ_e , is $1100 \text{ cm}^2/\text{V}\cdot\text{sec}$, while the mobility of holes, μ_h , is $100 \text{ cm}^2/\text{V}\cdot\text{sec}$ (Sakai, 1982).

5.4 Transit time of electrons and holes in the detector

The time required for a charge carrier to travel from the site where it is created to the electrode at the edge of the crystal is called the transit time. Since the applied electric field is constant, the drift velocity in the detector is constant. If at time $t=0$ a gamma ray interacts with the crystal a distance b from the positive electrode, then the transit time for the electrons is:

$$t_e = \frac{b}{v_{de}} = \frac{b}{\mu_e E}, \quad (5.9)$$

and the transit time for the holes is:

$$t_h = \frac{L - b}{\mu_h E}, \quad (5.10)$$

where L is the distance between the electrodes.

5.5 The change in voltage at the detector electrodes

As soon as charge carriers created by a gamma-ray interaction begin moving through the semiconductor crystal a change in energy $\Delta\mathcal{E}$, occurs in the detector. To derive an analytic form for $\Delta\mathcal{E}$, consider the detector to be a capacitor with capacitance C . Then

the energy stored in the detector can be expressed as:

$$\mathcal{E}_{\text{stored}} = \frac{1}{2} C V^2 , \quad (5.11)$$

where V is the voltage difference between the electrodes of the detector. Now consider the electrons and holes created by the gamma-ray interaction to be two clouds of charge, q and $-q$. If the incident photon has energy $\mathcal{E}_\gamma$ then for CdTe:

$$q = \frac{e \mathcal{E}_\gamma}{4.43 \text{ eV}} . \quad (5.12)$$

The work done in moving a particle a distance x under a force $\mathbf{F}$ is:

$$\Delta \mathcal{E} = \int_0^x \mathbf{F} \cdot d\mathbf{x} = \int_0^x (q \mathbf{E}) \cdot d\mathbf{x} = q E x = q \frac{V}{L} x , \quad (5.13)$$

where $\Delta \mathcal{E}$ is the change in the energy of the detector.

Differentiating Equation 5.11 with respect to the voltage shows that a change in energy in the detector $\Delta \mathcal{E}$ causes a change in voltage ΔV :

$$\Delta \mathcal{E} = \frac{1}{2} C (2 V \Delta V) . \quad (5.14)$$

Equating Equations 5.13 and 5.14 yields the result:

$$\Delta V = \frac{q}{C L} x . \quad (5.15)$$

This equation predicts that the change in voltage at the detector electrodes increases linearly with a slope of q/CL as the distance between electron-hole pairs increases, reaching a maximum when the charge carriers reach the electrodes and are separated by the length of the crystal, L . However, an important mechanism has been left out of this calculation. As will be discussed below the process of trapping stops or delays charge carriers, so that the amount of charge in the detector varies nonlinearly with time.

To summarize the discussion so far, the method by which semiconductors detect gamma rays can be described in three steps: (1) A gamma ray interacts with the semiconductor detector producing tens of thousands of electron-hole pairs. (2) An electric field is maintained across the detector, and under its influence the holes and electrons migrate in opposite directions. (3) The charge separation between the holes and electrons produces a change in voltage at the detector electrodes. Since the average number of electron-hole pairs produced is proportional to the energy of the incident gamma ray, the magnitude of the change in voltage is proportional to the energy of the incident gamma ray.

5.6 Trapping effects

Semiconductor crystals do not have perfect lattice structures. Lattice defects can be caused by impurities in the material or structural defects from imperfect crystal growth or subsequent deformation. When a charge carrier encounters such a defect, it may become trapped there. This discussion of trapping follows the treatment of Mayer (1968).

Suppose the crystal defect creates an energy level in the forbidden zone between the valence and conduction bands. This trap can be filled by the capture of a conduction band electron or the promotion of a valence electron. Once in the trap the electron can be emitted into the conduction band through thermal excitation, or it can combine with a hole and drop into the valence band. At time t the average number of charge carriers is:

$$N(t) = N_0 e^{-t/\tau_f} , \quad (5.16)$$

where N_0 is the number of charge carriers at time $t=0$, and τ_f is the mean time that the charge carrier spends between traps, known as the mean free time. The electron mean free time, τ_e , depends on crystal purity and uniformity. For today's best CdTe τ_e is 100 nsec (Sakai, 1982). The effect of trapping is to reduce the total mobile charge in the detector. If at time $t=0$ a gamma ray interacts with the crystal a distance b from the positive electrode, then the change in voltage due to electrons is:

$$\Delta V(t) = \frac{q \mu_e \tau_e E}{L C} (1 - e^{-t/\tau_e}) \quad t \leq b/(\mu_e E) , \quad (5.17)$$

where C is the capacitance of the semiconductor crystal, and for the moment we are ignoring the possibility that an electron can escape from its trap (detrapping). The change in voltage due to holes is:

$$\Delta V(t) = \frac{q \mu_h \tau_h E}{L C} (1 - e^{-t/\tau_h}) \quad t \leq (L - b)/(\mu_h E) , \quad (5.18)$$

where again we are ignoring detrapping. Summing Equations 5.17 and 5.18 gives the total change in voltage in the detector as a function of time for times less than the transit times of the holes and electrons.

For times greater than the transit times, the voltage becomes constant:

$$V = \frac{q \lambda_e}{L C} (1 - e^{-b/\lambda_e}) + \frac{q \lambda_h}{L C} (1 - e^{-(L-b)/\lambda_h}) , \quad (5.19)$$

where $\lambda_e = \mu_e \tau_e E$ and $\lambda_h = \mu_h \tau_h E$.

Figure 5-2 is a graph of voltage versus time using Equations 5-17 through 5-19 for an interaction midway between the electrodes. The dashed line shows how detrapping would affect the voltage. Note that Equations 5-17 through 5-19 predict that the slope of the voltage pulse is dependent on where the interaction occurs in the crystal. If the interaction is near the negative electrode (large b), the holes will reach the negative electrode immediately, and the slope of the voltage pulse will be dominated by the electrons travelling toward the positive electrode. Since electron mobility is high, the slope of the voltage pulse will be steep. However, if the interaction is near the positive electrode (small b) the electrons will immediately reach the positive electrode and the slope of the voltage pulse will be dominated by the motion of holes toward the negative electrode. Since the mobility of holes in CdTe is an order of magnitude less than the mobility of electrons, the slope of the voltage pulse for an interaction at small b will be much less than the slope for an interaction at large b .

5.7 Pulse height spectrum

Figure 5-3 shows a typical ^{111}In pulse height spectrum. This spectrum was obtained by placing an ^{111}In source near a CdTe detector, amplifying the output from the detector and putting it into a multichannel analyzer or MCA. The MCA increments the channel number corresponding to the total energy of the pulse. The result is a histogram called a pulse height spectrum. The taller peak corresponds to the photopeak energy of the 171 keV gamma ray, while the shorter peak corresponds to the photopeak energy of the 245 keV gamma ray.

If all the photons incident on the detector had energies of either 171 keV or 245 keV, and the detector always produced an output pulse strictly proportional to the

Figure 5-2. Detector Voltage as a Function of Time. The solid line is a graph of Equations 5.17 through 5.19, showing the change in voltage at a detector electrode as a function of time, ignoring detrapping. The initial value, V_0 , is the voltage maintained across the detector. At $t=0$ a gamma ray interacts with the detector material midway between the electrodes. Between $t=0$ and $t=t_e$, the transit time of the electrons, the decrease in voltage reflects a combination of electrons moving toward the positive electrode and holes moving toward the negative electrode. At $t=t_e$ all electrons created by the gamma-ray interaction have reached the negative electrode. The voltage continues to decrease between $t=t_e$ and $t=t_h$ (though the slope is gentler) as the slower-moving holes continue travelling through the crystal. At $t=t_h$ the voltage in the detector becomes constant at the value given by Equation 5.19. The dotted line shows the change in voltage if detrapping is included. The voltage continues decreasing gradually as charge carriers are freed from their traps and move toward the electrodes. On a time scale much larger than that of this graph the voltage gradually returns to the initial value as the high-voltage power supply recharges the detector.

Figure 5-3. Typical ^{111}In Pulse Height Spectrum. This spectrum was generated by placing a point source of ^{111}In near a CdTe detector. The output from the detector was sent to a multichannel analyzer. For each input pulse, the MCA incremented a channel number corresponding to the total energy of the pulse. The taller peak in the spectrum corresponds to the photopeak energy of the 171 keV gamma ray of ^{111}In . The shorter peak corresponds to the photopeak energy of the 245 keV gamma ray. The 245 keV peak is shorter because the detector is less efficient in stopping higher energy gamma rays. The spread in energies about the photopeak is due to the Poisson distribution in the number of electron-hole pairs produced. The asymmetries toward the low energy side of the peaks are due mainly to charge carrier trapping. The shaded regions are typical choices for windows.

energy of the incident photon, the pulse height spectrum would be two delta functions located at 171 keV and 245 keV. Instead, there is a spread of energies about the photopeak energies and both peaks are asymmetric, having long tails on the low-energy side. This spread in energies about the photopeak arises three ways: (1) the incident photon can arrive at the detector with less than the photopeak energy, (2) the amount of charge produced by incident photons of the same energy can vary, and (3) even if the charge deposited in the detector by two photons is equal, the pulse height for the two interactions can vary. In the next three paragraphs we discuss each of these possibilities separately.

The spectrum in Figure 5-3 was obtained from a point source in air. When a semiconductor detector is operated as a surgical probe, however, there is tissue between the source and detector, and gamma rays can Compton scatter within the source volume before reaching the detector. Since Compton-scattered photons lose energy to the electrons they collide with, they arrive at the detector with less than the photopeak energy.

There are two reasons why the amount of charge produced in the detector can vary for incident photons of the same energy. The first is that the creation of electron-hole pairs is a random process governed by Poisson statistics, so the number of electron-hole pairs produced is a Poisson random variable. The second reason is that an incident photon can Compton scatter within the detector. If the scattered photon then escapes from the detector there is much less energy available to create electron-hole pairs and these events will appear far to the left in the spectrum.

Finally, we have already described two reasons why the pulse height can vary even when the amount of charge is the same. The first is that the shape of the voltage pulse depends on where the interaction takes place within the detector. Because the site of the interaction is random, the pulse height varies randomly. The second reason the pulse height can vary is trapping. Trapping is a random process, so the number of charge carriers that

are trapped (and hence removed from the pulse) is a random variable. Trapping is the major contributor to the long tails on the low-energy side of the photopeaks.

5.8 Windowing the pulse height spectrum

Compton scattering in the body is especially harmful to the quality of the object reconstruction. The reconstruction method assumes all photons travel straight paths to the detector so Compton-scattered photons will be incorrectly positioned in the reconstruction. Therefore we would like to exclude Compton-scattered photons from the data. It is relatively easy to exclude photons that Compton scatter in the detector because they appear at very low energies in the spectrum so a simple low energy threshold will work. It is much more difficult to exclude photons that Compton scatter in the object volume, because their position in the spectrum overlaps the position of nonscattered photons whose lower pulse height is due to trapping. Regardless of what threshold we select we will exclude nonscattered photons from the data along with the scattered photons, thus reducing the efficiency of the detector. The process of choosing thresholds about the photopeaks is called windowing. The windows indicated on Figure 5-3 are typical of the windows we selected.

5.9 Real world detectors versus ideal detectors

As the above discussion shows, the major effect of modelling detectors as being ideal, as we did in our computer simulations, is that we significantly overestimate the detector efficiency. For example, the CdTe detectors and window settings chosen for the laboratory experiments discussed in the next chapter produced efficiencies of about 34% for the 171 keV gamma ray of ^{111}In and only 5% for the 245 keV gamma ray. Detector efficiencies are particularly important in coincidence experiments because two gamma rays must be

detected for each coincidence event. Therefore the number of coincident photons detected is proportional to the square of the detector efficiency.

CHAPTER 6

LABORATORY EXPERIMENTS

The simulation results presented in Chapter 4 rely on the analytic expression used for the system matrix, H , given in Equation 2.7. We know that a real-world coincidence imaging system will differ from this expression. For instance, Equation 2.7 does not include the slight correlation between the two gamma rays of ^{111}In or the efficiency of the detectors or the effect of attenuation and scatter. If the real-world system matrix for coincidence imaging differs greatly from the analytic expression, then predictions of probe performance based on the simulations are invalid. In this chapter we describe laboratory experiments conducted to verify the general form of the system matrix given by Equation 2.7. We constructed a circuit to measure coincidences between two CdTe detectors and took measurements to try to determine whether the coincidence count rate followed the relationship predicted in Equation 2.7 when (1) the detector-detector separation varied, and (2) the detector-source separation varied. One important feature to confirm was that the coincidence rate falls off as $1/z^4$, since that characteristic is the key to the prediction that a collimatorless coincidence probe can suppress contributions from the distant background. A second important feature to verify was that the measured ratio between true coincidences and accidental coincidences agreed well with the expected ratio, to confirm that the coincidence circuit was operating properly.

We simulated the 10×10 array by measuring coincidences between two CdTe detectors. Figure 6-1 is a closeup photograph showing one of these $0.2\text{cm} \times 0.2\text{cm} \times 0.2\text{cm}$ detectors with electrodes attached to opposite sides. The two detectors were placed on a $2\text{cm} \times 2\text{cm}$ grid, representing the proposed 10×10 array. By moving the detectors with respect to each other, we investigated the effect of varying the detector spacing. By moving

Figure 6-1. Closeup Photograph of Detector in Laboratory Setup. The black cube slightly below the center of the photo is one of the two $0.2\text{cm} \times 0.2\text{cm} \times 0.2\text{cm}$ CdTe detectors used in the experiment. The detector is epoxied to a rectangle of plexiglas or "boat" which is 0.5 cm long, 0.2 cm wide and 0.05 cm thick. The positive electrode of the detector sits in a circular depression in the plexiglas (visible on the second, empty boat in the photo). Behind the detector is the $2\text{cm} \times 2\text{cm}$ grid we used to simulate the 10×10 detector array. The detector/boat combination is held to the grid by double-stick tape. To the right of the detector is the threaded shank containing a small volume "point" source. The source was constructed by soaking a piece of paper in a solution containing ^{111}In , then drying it and ramming it into the hollowed-out tip of a threaded shank. The shank was mounted on two micrometer stages, allowing horizontal movement of the source parallel to the plane of the detectors (labelled the x direction) and perpendicular to that plane (labelled the z direction). The detectors, source and micrometer stages were placed in an aluminum box with a tightly fitting lid to shield the detectors from room lights and electromagnetic fields.

the source parallel and perpendicular to the plane of the detectors, we explored the effect of varying the detector-source distance. Finally, by leaving the detectors stationary and letting the source decay, we confirmed the predicted relationship between true and accidental coincidence rates and the source activity.

6.1 Laboratory setup

Figure 6-2 is a schematic of the coincidence circuit built to operate the two detectors. The two semiconductor detectors used to count coincident events were labelled A and B. The circuit restricts coincident events to those where gamma rays of different energies (171 keV and 245 keV) interact in different detectors within a coincidence interval of 4 μ sec. A true coincidence occurs when those photons come from the same decay. An accidental coincidence results if the photons are from different decays. Requiring that the photons have different energies eliminates half of the possible accidental coincidences. However this circuit cannot recognize coincidences, either true or accidental, that occur when two photons strike the same detector.

At bottom left in Figure 6-2 is one of the CdTe detectors. One of its electrodes is grounded, while the other is connected through a load resistor to a high voltage power supply. When a gamma ray interacts with the detector material, the detector produces a pulse. Chapter 5 contains details of the operation of semiconductor detectors.

The CdTe detector is capacitively coupled to a charge-sensitive integrating pre-amplifier, consisting of an operational amplifier and feedback capacitor C_2 and resistor R_2 . Operational amplifiers act to maintain equal voltages at their inputs, so when a voltage difference from a gamma-ray interaction appears at the ungrounded input, capacitor C_2 charges to drop that input back to ground. The resistor and capacitor combination $R_2 C_2$ is chosen so that it takes 50 μ sec for the voltage at the operational amplifier output to return to

Figure 6-2. Diagram of Coincidence Circuit. Coincidence is measured between two CdTe detectors, represented by squares labelled A and B on the left side of the schematic. The "wavy" arrows on the left of the detectors represent incoming gamma rays. One electrode of each detector is grounded, while the other is connected through a load resistor (R_L) to a high voltage power supply (HV). Each detector is connected to a charge-sensitive preamplifier, consisting of a capacitor C_1 and an operational amplifier. The output of the operational amplifier is sent to a linear amplifier, or shaping amplifier, labelled SA in the diagram. The output from the shaping amplifier is split and sent to two single-channel analyzers (SCAs). For each detector, one SCA is set to output a logic pulse when it receives a pulse corresponding to a 171 keV gamma ray interacting in the detector, while the other SCA is set to output a logic pulse when it receives a pulse corresponding to a 245 keV gamma ray interacting in the detector. The logic pulses generated by the SCAs are sent to one shots that lengthen the logic pulses, then to two AND gates that act as coincidence units. One AND gate receives the logic pulses corresponding to 171 keV gamma rays interacting in detector A (which we labelled Channel A1) and pulses corresponding to 245 keV gamma rays interacting in detector B, labelled Channel B2. The second AND gate receives 171 keV pulses from detector B, labelled Channel B1, and 245 keV pulses from detector A, labelled Channel A2. The AND gate inputs and outputs were sent to counters on a data acquisition board installed in a personal computer.

zero. Gamma ray interactions can occur more frequently than that, so pulses from different interactions overlap and sum, leading to a jagged, sawtooth-shaped signal. The output of the operational amplifier is sent to a linear, shaping amplifier. In addition to amplifying the signal, the shaping amplifier differentiates it so that each upward step in the signal corresponding to another gamma ray interaction becomes a narrow output pulse with a pulse height proportional to the energy deposited by the gamma-ray interaction.

The output from the shaping amplifier is split and sent to two single-channel analyzers (SCAs). Each SCA compares the input pulse height to two user-chosen voltage levels, a lower threshold and an upper threshold. If the pulse height is between these two levels a TTL logic pulse is generated. In our experiment, one SCA was set to produce a logic pulse if the incoming pulse had an amplitude corresponding to the photopeak energy of the 171 keV gamma ray of ^{111}In , while the second was set to logic output a pulse when the incoming pulse had an amplitude corresponding to the photopeak energy of the 245 keV gamma ray. As discussed in Chapter 5 there is a spread in the energies of the pulses from the detector. Therefore a range of energies below the photopeak energy must be accepted or the detector efficiency will be unacceptably low. Figure 5-3 shows typical choices for these windows.

The complete circuit had two detectors, two preamplifiers, two shaping amplifiers, and four SCAs. In our original circuit, the logic pulses from the SCAs were sent directly to two AND gates that acted as coincidence units. One AND gate received the logic pulses corresponding to 171 keV gamma rays interacting in detector A (which we labelled Channel A1) and pulses corresponding to 245 keV gamma rays interacting in detector B, labelled Channel B2. The second AND gate received 171 keV pulses from detector B, labelled Channel B1, and 245 keV pulses from detector A, labelled Channel A2. The level diagram for ^{111}In (Figure 2-2) shows that the half-life of the 245 keV ^{111}Cd atom is

85 nsec, indicating that the mean delay between the emission of the 171 keV and 245 keV photons is $0.12 \mu\text{sec}$ ($=85 \text{ nsec}/\ln 2$). The logic pulses from the SCAs are five times longer than that, $0.6 \mu\text{sec}$, which we thought would ensure that when a coincident event occurred the pulses would overlap at the AND gate and an output pulse would be generated.

The AND gate inputs and outputs were sent to counters on a data acquisition board installed in a personal computer. After being initialized to zero, the counters accumulated counts for a user-specified number of seconds, and the results were written to disk. The data sets collected consisted of six columns of integers, representing the accumulated counts in the four inputs to the AND gates: 1) channel A1, 2) channel A2, 3) channel B1, 4) channel B2; and the two outputs of the AND gates: 5) channel A1 AND channel B2, and 6) channel A2 AND channel B1. The sum of the last two columns is the total number of coincidence events recorded. The counts in individual channels will be used later to estimate the rate of accidental coincidences in the data. Figure 6-3 is a photograph showing much of our laboratory setup.

One circuit element in Figure 6-2 has not been explained, the one-shots between the SCAs and AND gates. The next section describes how we came to realize the necessity of including one-shots in our circuit.

6.2 Determining that the circuit measures true coincidences

Since the circuit cannot distinguish between true and accidental coincidences, both are included in the total number of coincidences recorded. It is not possible to separate out true coincidences, but it is possible to estimate the accidental coincidences. As discussed above, the mean time between the emission of the 171 and 245 keV photons is about $0.12 \mu\text{s}$. Suppose we delay the logic pulses going into one of inputs of the AND gate by a relatively long time, e.g. $16 \mu\text{sec}$. Since the pulses are only $0.6 \mu\text{sec}$ long, this long delay

Figure 6-3. Photograph of Laboratory Setup. In the right background is a Nuclear Instrumentation Module (NIM) bin containing the high-voltage power supply, shaping amplifiers, and single channel analyzers. In the foreground are the two charge-sensitive preamplifiers. In operation the preamplifiers would be connected to an aluminum box containing the detectors and source (not shown). Hidden from view in the rear are the AND gates and one shots and the cable connecting the circuit to a data analysis board in a personal computer. On the left is a multichannel analyzer which we used to determine the correct window settings for the SCAs.

would ensure that pulses from true coincidences never overlapped at the AND gate, so no output pulse would be produced. Accidental coincidences would continue to be counted, however, since their rate depends only on the length of the coincidence interval and not on the timing between channels.

We performed this experiment by using the logic pulses from one SCA as the external trigger to a pulse generator. We set the output pulse to be the same magnitude and length as the input pulse but delayed in time. As the delay was increased, we expected to see the number of coincidences drop abruptly at the point where the pulses from true coincidences no longer overlapped and we were measuring only accidental coincidences. Instead the number of coincidences remained almost constant, regardless of the amount of delay introduced. We calculated the number of accidental and true coincidences to expect according to Equations 2.8 and 2.11 and found that the number of coincidences we measured was consistent with measuring mostly accidentals. We had to determine why we were not measuring the majority of the true coincidences.

After some thought we realized that our circuit failed to take into account the transit time of electrons and holes through the crystal. In our experiment the applied voltage, V_0 , is 90 V, the length of the crystal, L , is 0.2 cm, and the mobility of the holes, μ_h , is $100 \text{ cm}^2/\text{V}\cdot\text{sec}$. If we plug these values into Equation 5.7 we see that holes created near the positive electrode that must travel the length of the crystal to reach the negative electrode could take as long as $4.4 \mu\text{sec}$ to arrive. Therefore the difference between the onset of the pulse at the SCA and its peak could be as much as $4.4 \mu\text{sec}$. As described in Chapter 5 the mobility of electrons is an order of magnitude greater than the mobility of holes, so the drift velocity of electrons is an order of magnitude larger than that of holes. We will concentrate on holes for the remainder of this discussion because they dominate the transit time effect. The problem arises because the SCA produces a TTL pulse only after

the peak of the incoming pulse is detected. Suppose the other gamma ray in a coincidence event interacts close to the negative electrode. Those holes would have only a short distance to travel so the peak of that pulse would arrive rapidly at its SCA, and a TTL pulse would be generated immediately. In this situation the difference in the arrival times of the two pulses at the AND gate could approach $4.4 \mu\text{sec}$. Since the TTL pulses are only $0.6 \mu\text{sec}$ long, the pulses would not overlap at the AND gate and the coincidences would not be counted.

There are two ways to ensure that true coincidence events are detected. The first is to replace the SCAs with circuits that produce logic pulses at the onset of the pulse, instead of at its peak. This kind of circuit is called a fast/slow coincidence circuit. The fast part is the timing circuit that determines whether a coincidence has occurred and triggers at the beginning of the pulse from the shaping amplifier. The slow part is the determination of the energy of the pulse which must wait for the peak value to arrive. The second alternative is to lengthen the $0.6 \mu\text{sec}$ TTL pulses until pulses from true coincidences overlap at the AND gate regardless of where the gamma ray interacts in the detector. We adopted the second alternative since it was simpler and cheaper, and added one-shots to increase the pulse lengths before the AND gates.

6.3 Deciding how much to lengthen the TTL pulse

A review of Equations 2.8 and 2.11 leads us to expect that as we lengthen the coincidence interval, δt , the ratio of true coincidences to accidental coincidences will increase until all true coincidences are being measured. Then as we further lengthen δt the ratio will decrease, since the number of true coincidences will remain constant while the number of accidental coincidences will continue to increase. To determine what coincidence interval maximized this ratio, we tested the system at three pulse lengths: $\delta t = 4.0 \mu\text{sec}$,

4.5 μsec and 5.0 μsec . We collected data first with zero delay between the channels to obtain measured total (true plus accidental) coincidence rates, $R_{\text{tot}}^{(m)}$. Then we collected data with a 16 μsec delay between channels to obtain measured accidental coincidence rates, $R_a^{(m)}$. We calculated the measured true coincidence rate by taking the measured total coincidence rate minus the measured accidental rate:

$$R_t^{(m)} = R_{\text{tot}}^{(m)} - R_a^{(m)} . \quad (6.1)$$

Then we calculated the ratio of **measured** true coincidences, $R_t^{(m)}$, to **measured** accidental coincidences, $R_a^{(m)}$ for A1 AND B2 and A2 AND B1:

$$\rho^m = \left(\frac{R_t^{(m)}}{R_a^{(m)}} \right) . \quad (6.2)$$

This ratio is listed in Table 6-1.

From Equation 2.12 we know that the ratio of the **expected** true coincidences to the **expected** accidental coincidences for a point source is:

$$\rho^e = \frac{R_t^{(e)}}{R_a^{(e)}} = \frac{1}{A \delta t} . \quad (6.3)$$

This ratio is also listed in Table 6-1.

To compare the results for the three pulse lengths we normalized the measured ratios by dividing by the expected ratios:

Table 6-1
Effect of Varying Length of Coincidence Interval

Pulse	$R_t^{(m)} / R_a^{(m)}$		$R_t^{(e)} / R_a^{(e)}$	ρ^m / ρ^e	
	A1xB2	A2xB1		A1xB2	A2xB1
4.0 μs	2.00	2.23	2.71	0.738	0.823
4.5 μs	1.55	1.93	1.93	0.805	1.00
5.0 μs	1.23	1.38	2.09	0.589	0.661

$$\frac{\rho_m}{\rho_e} = \frac{\left[R_{\text{tot}}^{(m)} - R_a^{(m)} \right] / \left[R_a^{(m)} \right]}{1/(A \delta t)} . \quad (6.4)$$

The result is the two rightmost columns of Table 6-1. From a pulse length of 4.0 μsec to a pulse of 4.5 μsec the ratio increased, indicating that a coincidence interval of 4.0 μsec did not count all of the true coincidences. From pulse lengths of 4.5 μsec to 5.0 μsec the ratio decreased, indicating that the number of true coincidences had stabilized while the number of accidental coincidences continued to rise. Therefore, the optimum coincidence interval was 4.5 μsec .

6.4 Placement of detectors and description of grid coordinate system

In Figure 6-1 the 2cmx2cm grid representing the full 10x10 array is visible behind the detectors. The top left detector position was labelled $(y,x) = (1,1)$ and the bottom right position was labelled $(10,10)$. The source was fixed vertically at grid position $y = 5.25$. Because accurate repositioning of the detectors on the grid was difficult, the detectors were left in the same positions for most of the data-taking sessions, as close as possible to $(y,x) = (5,5)$ and $(5,6)$.

6.5 Determining the coordinates of the detectors -- scans in x

To determine the micrometer reading corresponding to the centers of the detectors, we performed a lateral scan across the detectors (a scan in x). At each position ten 20-second integrations were collected. The initial x position of the source was about halfway between the detectors, at micrometer reading 1.511 cm. The source was stepped in 0.064 cm increments to the left (towards detector A) until the source was well past the detector. Then repeat readings were taken at positions 1.257 and 1.511 cm. Next the source

was stepped in 0.064 cm increments to the right (towards detector B) until the source was well past detector B. Repeat readings were taken at 1.765 cm and 1.511 cm. Figure 6-4 is a graph of the first and third columns of this data set, corresponding to counts recorded in channels A1 and B1. The solid lines are best-fit polynomials. The left-hand peak is the center of detector A. For simplicity, we decided to arbitrarily equate the micrometer reading at the center of detector A, $x = 1.328$ cm to grid coordinate $(y, x) = (5, 5)$. The right-hand peak is the center of detector B and is located at $x = 1.631$ cm. By recalling that the grid spacing is 0.2 cm, we can derive the grid coordinate for the center of detector B, $(y, x) = (5, 6.51)$.

We can derive the efficiency of detectors A and B by comparing the measured counts in channels A1 and B1 to the expected number of counts obtained by multiplying Equation 2.1 by Equation 2.11:

$$N_{\text{single phtns}}^{(e)} = N_{171} P(1 \text{ phtn}) = 0.902 A \Delta t \frac{\epsilon^2}{4\pi} \frac{\cos^3 \theta_1(R)}{z^2}. \quad (6.5)$$

For this experiment the activity of the source, A , was $2.59 \mu\text{Ci}$ ($= 9.58 \times 10^4 \text{ Bq}$), the counting time, Δt , was 20 seconds, and the area of the detector, ϵ^2 , was 0.04 cm^2 . The efficiency of both detectors was about 36% for the 171 keV gamma ray. For the 245 keV the efficiency of detector A was about 5.8%, while the efficiency of Detector B was 5.0%.

6.6 Scans in z

We also collected data in the z direction. The source's position in x was kept approximately midway between the detectors, at micrometer reading 1.511 cm. When the micrometer in the z direction read zero, the distance from the center of the source to the plane containing the detector centers was $0.25 \text{ cm} \pm 0.05 \text{ cm}$ (measured with a ruler).

Figure 6-4. Graph of Counts in Channels A1 and B1 during x scan. The squares are counts recorded in Channel A1, while the triangles are counts recorded in Channel B1. The source was $0.38 \text{ cm} \pm 0.05 \text{ cm}$ from the detector plane. Ten 20-second-long readings were taken at each position. The initial x position of the source was micrometer reading 1.511 cm, about halfway between the detectors. The source was stepped in 0.064 cm increments to the left (towards detector A) until micrometer position 1.003 cm. Then repeat readings were taken at positions 1.257 and 1.511 cm. Next the source was stepped in 0.064 cm increments to the right (towards detector B) until micrometer position 2.019 cm. Then repeat readings were taken at positions 1.765 and 1.511 cm. The estimated error in each micrometer reading was 0.005 cm. The solid lines are best-fit polynomials. The center of the peak on the left, $x=1.328 \text{ cm}$, is the center of detector A (marked by a large x), while the center of the peak on the right, $x = 1.631 \text{ cm}$, is the center of detector B.

Therefore to convert z micrometer readings to distance from the detector plane, we added a 0.25 cm offset.

The initial z position of the source was as close as possible to the detectors without touching them, at micrometer position 0.025 cm. From this initial position the source was stepped out from the detectors in increments of 0.064 cm to a position $z = 0.762$ cm from the detector plane. Then the source was stepped in larger increments of 0.254 cm to the limit of the micrometer, at $z = 2.79$ cm from the detector plane. However, if the source was weak enough that the collection time required to get coincident events at a micrometer setting exceeded 60 seconds, the source was not moved any farther away. At each position 10 readings were collected. However, unlike the x scans, the length of the collection time varied, ranging from 1 second to 60 seconds. To check the stability of the system, repeat readings were taken at the end of each run, usually at $z = 1.012, 0.758, 0.504$ and 0.282 cm. Table 6-2 contains a chronologic list of all the data sets used in the data analysis. The column of activities shows that four different sources were used.

6.7 Data reduction

A reference time at the beginning of the run was chosen, and the source strength at the reference time was calculated according to the formula:

$$A = A_0 2^{-\frac{\Delta t}{t_{1/2}}}, \quad (6.6)$$

where A_0 is the initial source strength (measured in a well counter), Δt is the elapsed time since the initial source strength measurement, and $t_{1/2}$ is the half-life of ^{111}In ($= 2.807$ days). The same formula was used to convert all subsequent measurements in the run to the same reference time.

Table 6-2
List of Data Used in Data Analysis

Type of data set	Activity in μCi
delay between channels = 4.5 μsec	3.12
delay between channels = 5.0 μsec	2.59
scan in x (lateral)	2.59
delay between channels = 4.0 μsec	2.49
scan in z (depth)	2.46
scan in z (depth)	1.86
scan in z (depth)	1.24
scan in z (depth)	0.72
scan in z (depth)	0.44
scan in z (depth)	0.16
scan in z (depth)	22.92
scan in z (depth)	8.81
scan in z with attenuator	8.81
scan in z (depth)	8.81
scan in z (depth)	4.11
scan in z (depth)	13.68
scan in z with attenuator	13.68
scan in z (depth)	13.68
scan in x (lateral)	20.55
scan in z (depth)	20.55
scan in x (lateral)	20.55
scan in z (depth)	20.55
scan in z (depth)	15.89
scan in z (depth)	12.29
scan in z (depth)	10.00
scan in z (depth)	7.73
scan in z (depth)	4.77

The measured coincidence rate is the sum of the coincidence rates of A1 AND B2 and A2 AND B1:

$$R_{\text{tot}}^{(m)} = R_{A1B2} + R_{A2B1} . \quad (6.7)$$

Therefore at each micrometer position we calculated the mean and standard deviation of the ten readings A1 AND B2 and A2 AND B1, converted the means and standard deviations to the reference time according to Equation 6.6 and summed the result.

The measured coincidence rate contains both true and accidental coincidences. According to Equations 2.9 and 2.13, both the number of true coincidences and the number of accidental coincidences is proportional to $(\cos^3\theta_i(\mathbf{R}) \cos^3\theta_j(\mathbf{R}))/z^4$, where θ_i is the angle between the normal to the i^{th} detector and the line connecting the i^{th} detector to location $\mathbf{R}$, θ_j is the angle between the normal to the j^{th} detector and the line connecting the j^{th} detector to location $\mathbf{R}$, and z is the perpendicular distance to the detector plane. Let us substitute z/D_A for $\cos\theta_i$ and z/D_B for $\cos\theta_j$, where D_A is the distance from location $\mathbf{R}$ to detector i and D_B is the distance from $\mathbf{R}$ to detector j . Then the number of true and accidental coincidences is linearly proportional to the quantity $z^2/(D_A^3 D_B^3)$. Therefore if we plot the measured coincidence data versus this quantity we should see a linear trend. Figure 6-5 shows a typical result. The line is the best fit in the least squares sense. In general the linear fit to the data was quite good.

It is possible to use the rest of our data, the count rates in the individual channels, to calculate the rate of accidental coincidences, $R_a^{(m)}$. If the rates in the individual channels are labelled R_{A1} , R_{A2} , R_{B1} and R_{B2} , then:

Figure 6-5. Graph of Total Coincidence Count Rate as a Function of $z^2/(D_A^3 D_B^3)$. As explained in the text a plot of the total measured coincidence count rate (true plus accidental coincidences) versus the quantity $z^2/(D_A^3 D_B^3)$ should exhibit a linear trend. The quantity z is the vertical distance from the source to the detector plane and D_A and D_B are the distances from the source to detectors A and B respectively. Note that because the abscissa has units of cm^{-4} , the higher the x value, the closer the source is to the detectors. The error bars are the standard deviation of the data measurements at each source position. The line is the best fit to the data in the least squares sense. The equation for the line is $y = 2.61 \times 10^{-2} x - 1.56 \times 10^{-2}$. The sum of squares of the residuals is 3.66×10^{-2} .

$$R_a^{(m)} = \delta t(R_{A1}R_{B2} + R_{A2}R_{B1}), \quad (6.8)$$

where δt is the length of the coincidence interval. Then the rate of true coincidences is the difference between Equations 6.7 and 6.8:

$$R_t^{(m)} = R_{\text{tot}}^{(m)} - R_a^{(m)} = R_{A1}R_{B2} + R_{A2}R_{B1} - \delta t(R_{A1}R_{B2} + R_{A2}R_{B1}). \quad (6.9)$$

Figure 6-6 shows the result of the true coincidence calculation on the same data set plotted in Figure 6-5. Again the abscissa is the quantity $z^2/(D_A^3 D_B^3)$, and a best-fit line is plotted. The same description applies to this type of graph as to Figure 6-5: the linear fit was in general quite good.

6.8 The variation in accidental and true coincidence rates with activity

According to Equation 2.8 the number of true coincidences should vary linearly with the activity, while Equation 2.11 indicates that the number of accidental coincidences should vary as the square of the activity. The sum of the two rates, the measured coincidence rate, should therefore follow a parabola when plotted against activity. To test whether our data followed this pattern, we plotted the slopes of the best fit lines from the measured coincidence data sets against activity. Figure 6-7 shows the initial result. The data points from the first two ^{111}In sources fell on the same parabola, while the data points from the second two sources fell on different parabolas. The difference originated in the way we made the different sources. Recall that we made sources by repeatedly dripping a liquid containing ^{111}In onto a small piece of paper, letting it dry, then dripping more liquid onto it until the paper contained the desired activity. The paper was then stuffed into the hole at the end of the threaded shank. When the first two sources were made, the piece of

Figure 6-6. Graph of True Coincidence Count Rate as a Function of $z^2/(D_A^3 D_B^3)$. The ordinate is the true coincidence count rate, obtained by subtracting the accidental coincidence count rate from the total coincidence count rate using Equation 6.9. The abscissa is the quantity $z^2/(D_A^3 D_B^3)$ defined in Figure 6-5. Note that because the abscissa has units of cm^{-4} , the higher the x value, the closer the source is to the detectors. As explained in the text the graph should exhibit a linear trend. The line is the best fit to the data in the least squares sense. The equation of the line is $y = 2.27 \times 10^{-2} x - 8.44 \times 10^{-3}$. The sum of squares of the residuals is 3.05×10^{-2} . The error bars were generated by a Gaussian propagation of errors calculation on Equation 6.9.

Figure 6-7. Plot of Slopes of Best-Fit Lines from the Total Coincidence Data Sets vs. Activity without Distance Correction. To gauge whether the measured true and accidental coincidence rates followed the predicted relationship, the detectors were left in the same position and measurements were taken as the source decayed. A graph of the form described in Figure 6-5 was generated for each data set and a best-fit line for the total coincidence count rate was calculated. Here the slopes of the best-fit lines are plotted versus the source activity. As explained in the text, the plot should follow a parabola. Although the data points from the first two ^{111}In sources do fall on the same parabola, the data points from the second two sources do not. The equation of the solid parabola is $y = 2.00 \times 10^{-3} x^2 + 1.36 \times 10^{-2} x + 4.22 \times 10^{-3}$. The sum of squares of the residuals is 3.85×10^{-4} . The equation of the upper dashed parabola is $y = 5.70 \times 10^{-3} x^2 + 8.28 \times 10^{-4} x + 1.31 \times 10^{-1}$. The sum of squares of the residuals is 2.24×10^{-3} .

paper used was small enough to allow a second untreated piece of paper to be pushed into the hole after the first. The second piece of paper acted as a stopper to ensure that none of the treated paper fell out. However, when the second two sources were made, the treated pieces of paper were larger, and there was no room to insert a second, untreated piece. The result was that the source was closer to the detector plane, and the conversion factor we used in the data reduction program, 0.25 cm, was incorrect. If we postulate that the difference in position was 0.07 cm, and recalculate the measured coincidence rates of the affected data runs using a conversion factor of 0.18 cm instead, we generate Figure 6-8. All of the slopes fall along the same parabola, confirming the expected relationship between the number of coincidences and the activity.

6.9 Effect of an attenuating medium

The experiment described so far was conducted with the source in air. A probe in the body, however, encounters the attenuating effects of tissue. To test how this might affect a collimatorless coincidence probe, we inserted a rectangle of plexiglas 5.0cmx6.6cmx3.8cm thick between the source and detectors. Plexiglas has approximately the same attenuation coefficient as water, so it is a reasonable approximation for human tissue. A hole was drilled in the plexiglas, through which the source could move. Measurements were taken with the source flush with the surface of the plexiglas, then at increasing depths within the plexiglas. Figure 6-9 is a graph of the best-fit lines for the true coincidence rates (bottom 3 lines) and measured total coincidence rates (top 3 lines) for three data runs on the same day. The dashed lines are the best-fit lines to the data sets with the attenuator. The percentage reduction in the number of coincidences caused by the attenuator is the same for both total and true coincidences. Combining these data sets with a similar sequence taken at an activity of 8.81 μCi , we find from the slopes of the best-fit lines that

Figure 6-8. Plot of Slopes of Best Fit Lines from the Total Coincidence Data Sets vs. Activity with Distance Correction. We postulated that ramming an second, untreated piece of paper into the threaded shank after the paper treated with ^{111}In pushed the source back 0.07 cm. The second two ^{111}In sources had no untreated paper inserted. Therefore we recalculated the measured total coincidence count rates for those data runs using a source position 0.07 cm closer to the detector plane. With this distance correction, the slopes of the best-fit lines for all four sources fall on the same parabola. The solid parabola is the best-fit in the least squares sense to all four data sets. The equation of the solid parabola is $y = 2.12 \times 10^{-3} x^2 + 1.11 \times 10^{-2} x + 7.50 \times 10^{-3}$. The sum of squares of the residuals is 4.50×10^{-3} . For reference the dashed parabola is the best-fit to the first two ^{111}In sources, and its parameters are given in the caption of Figure 6-7. Notice how well the data points for the second two ^{111}In sources follow this parabola when the distance correction is applied.

Figure 6-9. Graph of True and Total Coincidence Count Rates for Data Runs with and without Attenuator. The graph shows three data sets taken on the same day: the first with the source in air, the second with a 5.0cmx6.6cmx3.8cm piece of plexiglas inserted between the source and detector plane, and a third, abbreviated data set with the source in air to check for drift in the system. The solid lines show the coincidence rates for data sets conducted with the source in air. The dashed lines show the coincidence rates when the plexiglas was inserted. The top three lines are the best-fit lines to the measured **total** coincidence count rates. The bottom three lines are the best-fit lines to the **true** coincidence count rates. The percentage reduction in the number of coincidences caused by the attenuator is the same for both total and true coincidences.

the attenuator causes a $5.51 \pm 0.16\%$ reduction in both total and true coincidence rates. Therefore a simple scale factor can be used to predict the reduction in the number of coincidences caused by human tissue.

6.10 Results

Our experiments measuring coincidence between two CdTe detectors: (1) confirmed the $1/z^4$ dependence predicted in Chapter 2, (2) confirmed the expected activity versus activity squared relationship between the number of true coincidences and the number of accidental coincidences, and (3) determined that the effect of an attenuating medium can be expressed as a simple scale factor.

CHAPTER 7

SUMMARY AND SUGGESTIONS FOR FUTURE RESEARCH

Through computer simulations and laboratory experiments we have demonstrated that collimatorless coincidence imaging is a feasible approach for imaging objects close to the detectors and that it may prove useful in the task of tumor detection in nuclear medicine. This chapter will review the work done and suggest areas for future research.

7.1 Summary

In Chapter 2 we discussed previous work in the field of coincidence imaging and described our idea: that coincidence imaging using detectors without collimators can achieve useful imaging of objects near the detectors, while ameliorating the efficiency problems of earlier coincidence systems. We derived the expression for the matrix describing how points in the object space would contribute to measurements in images created by such a system and used it to predict that distant background sources would be suppressed.

In Chapter 3 we described the computer simulations of our proposed implementation of a coincidence imaging system, a 10x10 array of collimatorless detectors. The Monte Carlo routine we wrote included the effects of attenuation, accidental coincidences, Poisson noise, and the (slight) correlation in direction between the gamma rays of ^{111}In .

In Chapter 4 we introduced the method used to reconstruct estimates of the original objects from the simulated data generated in the simulations. We demonstrated that in a realistic simulation of the body the collimatorless coincidence system has a superior ability to suppress inhomogeneities in the background radiation when compared to an equivalent collimated probe. We also performed eigenanalysis to investigate what types of objects our detector array could, in principle, image perfectly (except for a scale factor).

In Chapter 5 we gave a brief description of the physics governing the behavior of semiconductor detectors.

In Chapter 6 we discussed laboratory experiments performed using two CdTe detectors. By moving a source vertically and horizontally with respect to the two detectors, and by moving the detectors with respect to each other we: (1) confirmed the $1/z^4$ dependence predicted in Chapter 2, (2) confirmed the expected activity versus activity squared relationship between the number of true coincidences and the number of accidental coincidences, (3) determined that the effect of an attenuating medium can be expressed as a simple scale factor, and (4) (not yet) confirmed that the effect of detector separation on coincidence rate follows the relationship predicted by the system matrix.

7.2 Suggestions for future research

Alternate geometries The coincidence configuration we chose, where coincidence occurs between equal-sized uncollimated detectors in a 10x10 detector array, was governed by a probe design our group already planned to build. Many alternative configurations for coincidence could be used and should be investigated. Figure 7-1 shows two possibilities. The top diagram is a system that would be used with a radionuclide that emits a low energy K x-ray and a higher energy gamma ray, such as ^{123}I . Coincidence occurs when the K x-ray interacts anywhere within the large detector on the bottom, while the gamma ray penetrates that detector and travels down a collimator bore to another detector. In the bottom diagram, coincidence occurs when one photon travels through a pinhole to an array of detectors, while the second photon interacts with a flanking detector. By changing the angle of the flanking detector, different regions within the object space could be explored.

Figure 7-1. Alternate Geometries for Coincidence Imaging. (a) The upper coincidence detector design is for use with a dual-photon emitting radionuclide in which one of the two photons has a much lower energy than the other. As indicated by the dashed lines, a coincidence event occurs when the lower-energy photon interacts with the slab detector while the higher-energy photon penetrates the slab detector and interacts with the collimated detector array behind. (b) The lower coincidence detector design consists of a detector array behind a pinhole (shielded by lead) with a flanking annular slab detector. As indicated by the dashed lines, a coincidence event occurs when one photon passes through the pinhole to the detector array behind, while the other photon interacts with the slab detector. This detector could be made sensitive to different depths in the object by varying the angle of the slab detector (as indicated by curved arrows).

Alternate methods of reconstruction While the SVD pseudoinverse reconstruction method we used gave encouraging results, it does not contain any prior information about the object space, such as the fact that values in real objects are positive. The possibility that reconstruction methods that do include prior information could achieve better reconstructions should be investigated. Our experimentation with the expectation-maximization (EM) algorithm without smoothing was too brief to be informative so it was not included in this dissertation.

Two suggestions for further avenues of research came from researchers in other groups:

Increase voxel size with depth Since there is a $1/z^4$ falloff in the number of coincidences with depth, it was suggested that the z dimension of the voxels in object space be scaled by z^4 for reconstructions. Then the probability of a coincidence event originating in a voxel would be more uniform throughout the object volume.

Comparison to noncoincidence collimatorless probe In comparisons between collimated conventional (noncoincidence) probes and our collimatorless coincidence probe, it is difficult to separate the contributions to image quality from (1) the suppression of distant background sources caused by the coincidence condition and (2) the possible increase in the number of photons reaching the detector plane due to the absence of collimators. Therefore it was suggested that it would be useful to run more simulations comparing our probe design to the same probe without coincidence.

REFERENCES

- Anger, H.O., (1958), "Scintillation camera", *Rev. Sci. Instrum.*, 29:27-33.
- Barrett, H.H., and W. Swindell, (1981), *Radiological Imaging: The Theory of Image Formation, Detection, and Processing*, Vols. I and II, Academic Press, Inc., New York, 693 p.
- Barrett, H.H., J.N. Aarsvold, and T.J. Roney, (1991), "Null functions and eigenfunctions: tools for the analysis of imaging systems," in *Information Processing in Medical Imaging*, D.A. Ortendahl and J. Llacer, eds.; Wiley-Liss, New York, pp. 211-226.
- Browne, E., and R.B. Firestone, (1986), *Table of Radioactive Isotopes*, John Wiley & Sons, Inc., New York, 1523 p.
- Chung, V., K.C. Chak, P. Zacuto, and H.E. Hart, (1980), "Multiple photon coincidence tomography," *Semin. Nucl. Med.*, 10(4):345-354.
- Hart, H.E., (1965), "Focusing collimator coincidence scanning," *Radiology*, 84:126.
- Hart, H.E., (1968), "Comparative resolution of single gamma counting and coincidence counting in focusing collimator scanning systems," *Trans. N.Y. Acad. Sci.*, 30:580-599.
- Hart, H.E., and S. Rudin, (1977), "Three-dimensional imaging of multimillimeter-sized cold lesions by focusing collimator coincidence scanning (FCCS)," *IEEE Trans. Biomed. Eng.*, BME-24:169-177.
- Hartsough, N.E., H.B. Barber, J.M. Woolfenden, H.H. Barrett, T.S. Hickernell, and D.P. Kwo, (1989), "Probes containing gamma radiation detectors for *in vivo* tumor detection and imaging," *Proc. SPIE*, 1068:92-99.
- Helmers, H., H.von Boetticher, and I. Schmitz-Feuerhake, (1979), "Scanner performance to overcome efficiency problems in three-dimensional scintigraphy by gamma-gamma-coincidences," *Phys. Med. Biol.*, 24:1025-1029.
- Helmers, H., H.von Boetticher, and I. Schmitz-Feuerhake, (1982), "Depth discrimination in direct 3D-scanning without image reconstruction using a coincidence technique," *Eur. J. Nucl. Med.*, 7:324-326.
- Hickernell, T.S., H.B. Barber, H.H. Barrett, and J.M. Woolfenden, (1988), "Dual-detector probe for surgical tumor staging," *J. Nucl. Med.*, 29:1101-1106.
- Hickernell, T.S., (1988), "Statistical decision-making with a dual-detector probe," Ph.D. Dissertation, University of Arizona.

- Hinman, G.W., G.R. Hoy, J.K. Lees, and J.C. Serro, (1964), "Angular correlation studies of electric field gradients in dilute silver alloys," *Phys. Rev.*, 135:A206-A217.
- Liang, Z., (1988), "Preliminary study of triple photon coincidence imaging technique," *Proc. SPIE*, 974:283-291.
- Liang, Z., H.E. Hart, and A. Schoenfeld, (1987), "Triple gamma coincidence tomographic imaging without image processing," in *Proceedings of the 9th Annual Conference of the Engineering in Medicine and Biology Society*, IEEE, New Jersey, pp. 825-826.
- Liang, Z., and R. Jaszczak, (1990), "Comparisons of multiple photon coincidence imaging techniques," *IEEE Trans. Nucl. Sci.*, 37:1282-1292.
- Mayer, J.W., (1968), "Search for semiconductor materials for gamma ray spectroscopy," in *Semiconductor Physics*, G. Bertolini and A. Coche, eds.; John Wiley & Sons, Inc., New York, pp. 445-498.
- Monahan, W.G., J.W. Beattie, and J.S. Laughlin, (1972), "Positron mode of the Total Organ Kinetic Imaging Monitor: system design and applications," *Phys. Med. Biol.*, 17:503-515.
- Morin, R.L., D.E. Raeside, J.E. Goin, and J.C. Widman, (1988), "Random number generation and testing," in *Monte Carlo Simulation in the Radiological Sciences*, R.L. Morin, ed.; CRC Press, Inc., Boca Raton, pp. 235.
- Packer, S., (1984), "Tumor detection with radiopharmaceuticals," *Semin. Nucl. Med.*, 14:21-30.
- Powell, M.D., (1989), "Multiphoton, time-of-flight three-dimensional radionuclide imaging," *Med. Phys.*, 16:809-812.
- Powell, M.D., W.G. Monahan, and J.W. Beattie, (1970), "Gamma-gamma coincidence detector," *Radiology*, 94:197.
- Press, W.H., B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, (1986), *Numerical Recipes*, Cambridge University Press, Cambridge, 818 p.
- Sakai, E., (1982), "Present status of room temperature semiconductor detectors," *Nucl. Instrum. Methods*, 196:121-130.
- Schmitz-Feuerhake, I., (1970), "Studies on three-dimensional scintigraphy with gamma-gamma-coincidence," *Phys. Med. Biol.*, 15:649-656.
- Sharp, P.F., P.P. Dendy, and W.I. Keyes, (1985), *Radionuclide Imaging Techniques*, Academic Press, Inc., Orlando, 271 p.

- Siffert, P., (1983), "Cadmium telluride detectors and applications," Mat. Res. Soc. Symp., in *Nuclear Radiation Detector Materials*, E.E. Haller, H.W. Kraner, and W.A. Higginbotham, eds.; North Holland: Elsevier Science Publishing Co., New York, pp. 87-114.
- Strang, G., (1980), *Linear Algebra and its Applications*, Academic Press, Inc., Orlando, 414 p.
- von Boetticher, H., H. Helmers, and Eckart-M. Muschol, (1979), "Contributions to depth discrimination gamma-gamma-coincidence methods in scintigraphy," Phys. Med. Biol., 24:571-576.
- von Boetticher, H., H. Helmers, P. Schreiber, and I. Schmitz-Feuerhake, (1982), "Advances in gamma-gamma-coincidence scintigraphy with the scintillation camera," Phys. Med. Biol., 27:1495-1506.
- Woolfenden, J.M., and H.B. Barber, (1989), "Radiation detector probes for tumor localization using tumor-seeking radioactive tracers," Am. J. Roentgenol., 153:35-39.
- Zubal, G., and C. Harrell, (1991), "Voxel based Monte Carlo calculations of nuclear medicine images and applied variance reduction techniques," in *Information Processing in Medical Imaging*, A.C.F. Colchester and D.J. Hawkes, eds.; Springer-Verlag, Berlin, pp. 23-33.