Title

Performance Characterization, Image Processing, and Multimodality Coregistration of Small Animal Positron Emission Tomography Systems

Permalink

https://escholarship.org/uc/item/0c11c24g

Author

Daver, Freddie

Publication Date

2012

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Performance Characterization, Image Processing, and Multimodality Coregistration of Small Animal Positron Emission Tomography Systems

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Biomedical Physics

by

Freddie Rohinton Daver

2012

ABSTRACT OF THE DISSERTATION

Performance Characterization, Image Processing, and Multimodality Coregistration of Small Animal Positron Emission Tomography Systems

by

Freddie Rohinton Daver

Doctor of Philosophy in Biomedical Physics

University of California, Los Angeles, 2012

Professor Magnus Dahlbom, Chair

Positron Emission Tomography (PET) systems designed specifically for small animal imaging have a unique set of requirements for successful operation. This work investigates a limited angle tomography microPET system, in contrast to conventional cylindrical systems. Initial studies included measurements of spatial resolution, uniformity, contrast-to-noise ratios, resolution recovery coefficients, and detector deadtime.

A subsequent investigation was performed to ascertain the use of Factor Analysis in mitigating artifacts caused by the limited angle tomography. Three separate studies were

conducted: a dual-isotope experiment, a semi quantitative study, and mouse models. The results indicate potential use of Factor Analysis with possible limitations due to noise or uptake.

A more rigorous investigation into the limits of Factor Analysis was performed using Principal Component Analysis. Using the definitions of correlation and covariance matrices, the results of uniform phantom studies were used to measure uniformity, signal-to-noise ratio, and noise distribution. This analysis was applied to the limited angle tomographic system and a conventional cylindrical system. The results were insightful and in strong agreement with Poisson statistics.

Multimodality coregistration between systems was also performed using a rigid body least-squares transformation ("Horn's method"). Spherical phantoms were imaged to obtain calibration data within each system. In addition to deriving the transformation between a pair of systems, a metric was computed to estimate the error associated with the coregistration. Once computed, the transformation was applied to independently acquired mice studies from each pair of systems. The method proved fairly successful, and a number of improvements are suggested.

The rigid body nature of the transformation also allowed for a convenient method of measuring spatial linearity. Coregistration between synthetic grid-like images, and actual images of grid-like phantoms, were performed. The least-squares metric produced is used to determine spatial linearity. Grid-like phantoms were created using activity-infused ink in conjunction with a conventional inkjet printer. Each phantom was scanned on the limited angle tomography system. The results indicated a very high degree of spatial linearity.

Although motivated by limited angle PET, most of the methods employed may be used for other imaging modalities as well. These methods are statistical in nature, and may be

similarly interpreted within various modalities. While the purpose of each study was different, the underlying mathematical concepts are common amongst most of them.

The dissertation of Freddie Rohinton Daver is approved.

Johannes Czernin

Henry Huang

Christiaan W.J. Schiepers

Michael Albert Thomas

Magnus Dahlbom, Committee Chair

University of California, Los Angeles

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

BIOGRAPHICAL SKETCH

University of California – Los Angeles

Graduate Student Researcher within the Biomedical Physics Interdepartmental Program

September 2006 – August 2012

Massachusetts Institute of Technology – Lincoln Laboratory

Full Time Staff Researcher

August 2000 – June 2006

Carnegie Mellon University

B.S. Physics

August 1996 – May 2000

1   Introduction

The following work consists of five separate studies. Although varied in topic, the motivation for each study is based on the limitations or conceptual similarities with the previous study. The first study consists of performance measurements taken with a limited angle tomography preclinical scanner for positron emission tomography (PET). These measurements include spatial resolution, uniformity, contrast-to-noise ratio (CNR), resolution recovery, and deadtime. Due to the geometrical configuration of the system, the resulting images suffer from artifacts that severely limit the sagittal and transaxial images both visually and quantitatively.

Factor Analysis (FA) was proposed in order to mitigate the detrimental effects of the limited angle tomography. Qualitatively, FA attempts to segment an image not by using spatial relationship between structures. Instead, FA attempts to group together voxels which exhibit similar kinetic properties over time. In other words, voxels which resemble each other, as measured temporally, are assumed to belong to the same structure. Using this technique, dynamically reconstructed images can be separated into factor images. Ideally, each factor image contains only one of the structures of interest. In order to study the applicability of FA on a limited tomography PET system, separate studies were performed. The first consisted of a dual-isotope measurement in order to determine how well FA could separate voxels from each of the two isotopes. The second study investigated the use of semi quantitative model in order to test the separation of overlapping simulated structures. The third study applied FA to a mouse scan in order to visually assess how successfully adjacent structures could be separated.

Only moderate success was achieved with FA. This motivated a simpler, more direct approach to quantifying how well FA might perform. This approach involved the use of Principal Component Analysis (PCA) which, qualitatively, may be thought of as a more restricted form of FA. For these studies, PCA was not used conventionally as a compression or image processing technique; instead, PCA was used on a set of uniform cylinder studies, of various duration and activity, in order to compute a number of performance metrics. These metrics include uniformity, a time-dependent signal-to-noise (SNR) ratio, noise distribution, as well as an upper bound limit for FA. Due to the novelty of this use of PCA, these methods were tested on the limited angle tomography system as well as a conventional cylindrical microPET system. In addition, due to the computational limitations of a conventional PCA calculation, an alternative method of computing the principal components was mathematically derived without the use of approximations, additional assumptions, or intermediate processing.

Aside from performance measurements and image processing techniques, multimodality coregistration between small animal imaging systems was investigated. Using methods adapted from PET brain motion correction, a coregistration method based on rigid body transformations was developed. The method is completely independent of modality, statistical properties, noise, or anatomical models. No approximations or iterative methods are used. Instead, the method relies almost entirely on how closely the assumption of rigid body is observed. In other words, motion of the subject relative to the fields-of-view (FOVs) is detrimental to the accuracy of this method. Measurements were performed in order to assess the positioning precision with which subjects could be placed within each FOV. Transformations between systems consisted of a rotation and translation. These transformations were computed using an analytic least-squares algorithm relying on previously acquired calibration data. A residual metric, indicating the

quality of these transformations, was also computed. Transformations were then used to coregister pairs of mice images acquired from pairs of modalities. Visual investigation of these mice images were then performed to give a qualitative evaluation of the coregistration.

The least-squares nature of the coregistration then motivated a test for spatial linearity. The discrepancies between a physical subject and its image may be described by a transformation. If this transformation is rigid, no spatial distortions are present. However, if spatial distortions are present, the relative location between any given points within the subject will differ from that in the image. Testing this spatial linearity was performed on the limited angle tomography system by coregistering a synthetic grid-like image with a physically acquired grid-like image. For both images, the grid spacing was identical. The coregistration was performed using the same rigid body transformation as that used for the multimodality coregistration. The resulting metric, provided by the transformation algorithm, is an absolute measure of the total mean squared error between grid points and their corresponding image, and, therefore, a measure of spatial linearity. The physical grids themselves were created by printing activity-infused ink onto high quality photographic paper. Coordinates of the synthetic grid points were computationally generated, while the imaged grid points were computed via centroid calculations. Both sets of coordinates were fed into the coregistration method in order to produce the spatial linearity metric.

In summary, drawbacks present in the limited angle tomographic system motivated a number of other investigations. Each investigation chronologically spawned subsequent investigations with unforeseen, but sometimes favorable, interpretations and consequences. Although seemingly disparate, the studies are strongly linked conceptually. Factor analysis and PCA are both methods of statistical separation. Furthermore, Principal Component Analysis,

rigid body coregistration, and spatial linearity all rely on the same fundamental least-squares concept. In fact, much of the work discussed in not specific only to PET, but other imaging modalities as well. As a result, a fairly comprehensive set of studies has been assembled.

2   Genisys System Characterization

2.1   Introduction

The Genisys system (Sofie Biosciences, Culver City, CA) is a Positron Emission Tomography (PET) system designed for preclinical studies [1, 2]. The low-cost Genisys system uses bismuth germanate (BGO) detectors arranged in a fixed dual head configuration, while most conventional microPET systems utilize lutetium oxyorthosilicate (LSO) detectors arranged in a cylindrical configuration. The dual head configuration of the Genisys is illustrated in Figure 2.1.
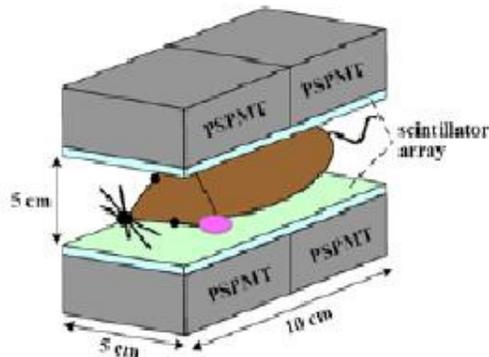
Figure 2.1: Dual head configuration of the Genisys microPET system. The lack of detectors in the sagittal direction gives rise to the image artifacts (i.e. "bowtie" artifact) which severely limits image quality. Due to the isotropic nature of radiation, detection must be performed with a full tomographic system in order to obtain statistically complete information.

The dual-head configuration of the Genisys reduces its geometrical efficiency and causes image reconstruction to be performed using limited angle tomography. Due to the strong dependence of PET images on geometrical factors, as well as on detector configurations, a comprehensive characterization of the Genisys system is needed in order to fully understand its characteristics and limitations.

2.2   Spatial Resolution Measurements

Spatial resolution can be defined as the minimum distance between two points at which they can be visually separated. The full-width at half-maximum (FWHM) is the most common method of quantifying spatial resolution. Due to geometrical effects, spatial resolution may be strongly dependent on position within the scanner field-of-view (FOV). Using a 50 µCi ^{22}Na point source, a 5-minute static scan was acquired at each location. The point source was repositioned from location to location with the aid of a high-precision translational stage. The locations of the point source were designed to fall on a grid within the central coronal plane of the Genisys. Distances between grid points were approximately 5 mm in the superior-inferior (SI) direction and approximately 1 mm in the left-right (LR) direction. Each image was reconstructed with the use of the Expectation Maximization (EM) algorithm with either 30 or 100 iterations. A Gaussian smoothing filter with a 0.35 mm standard deviation was also applied post-reconstruction. The raw images obtained from the system were composed of 80 x 80 x 176 voxels, each with dimensions of 0.55 mm x 0.55 mm x 0.55 mm. Figure 2.2 illustrates the complete set of all locations that were used.
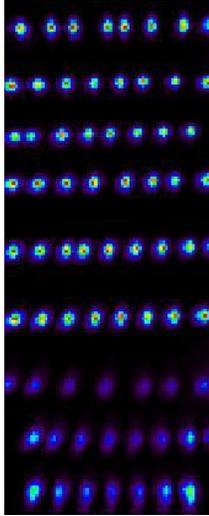
Figure 2.2: Coronal superposition of all images used in the spatial resolution measurements. The image show is 80 x 176 voxels (0.55 mm x 0.55 mm). Each individual image was obtained with the use of a 50 μCi ^{22}Na point source mounted on a translation stage. Each acquisition was five minutes in duration. The default 3D histogramming was used. Separate data sets used the default EM algorithm with 30 and 100 iterations. Spatial resolution was then found as prescribed by the NEMA protocol.

The calculation of spatial resolution was guided by the National Electrical Manufacturers Association (NEMA) Standards Publication NU 4-2008 [3]. All post-reconstruction analysis was performed on the raw Genisys images using the Interactive Data Language (IDL) version 7.0. The method consisted of adding the profiles of each point source image to obtain a "response function". The FWHM of the response function was then computed using a parabolic fit through its maximum. Points which corresponded to half of the maximum value of the parabolic fit were linearly interpolated. The reported FWHM is then defined as the distance between these interpolated points. Table 1, Table 2, and Table 3 indicate the computed FHWM values.

Direction	Full Width at Half of Maximum (mm)
Left-Right	1.32
Anterior-Posterior	2.82
Superior-Inferior	1.71

Table 1: FWHM of Genisys images using 30 iterations and no post-reconstruction smoothing filter.

Direction	Full Width at Half of Maximum (mm)
Left-Right	1.31
Anterior-Posterior	2.72
Superior-Inferior	1.60

Table 2: FWHM of Genisys images using 100 iterations and no post-reconstruction smoothing filter.

Direction	Full Width at Half of Maximum (mm)
Left-Right	1.58
Anterior-Posterior	3.01
Superior-Inferior	1.71

Table 3: FWHM of Genisys images using 100 iterations and a Gaussian post-reconstruction smoothing filter with 0.35 mm standard deviation.

For comparative purposes, the spatial resolution of the Siemens Inveon system is ~1.5 mm at FWHM [4, 5]. As expected from the detector configuration, the effects of limited angle tomography are most detrimental in the anterior-posterior (AP) direction. The post-reconstruction smooth filter is generally applied for to reduce noise while preserving spatial resolution and does not appear to have a significant effect. Measurements performed above or below the central coronal plane are expected to show increased FHWM values in all directions due to limited angle tomography as well as increased parallax effects.

2.3 Phantom Measurements

Due to the unconventional size and configuration of the Genisys system, an image quality phantom was constructed by Sofie Biosciences specifically for image quality control of the

Genisys [2]. This image quality phantom consists of two 1.0 mL cylindrical chambers; five rods with diameters of 1 mm, 2 mm, 3 mm, 4 mm, and 5 mm; and a rectangular pool of 14.2 mL. Both cylindrical chambers are engulfed within the pool. Note that the pool is filled with activity via injection into the rods; all rods and the pool share the same activity concentration when filled. A schematic diagram of the phantom is shown in Figure 2.3.
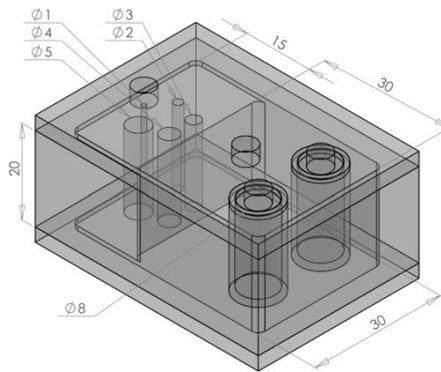
Figure 2.3: Schematic drawing of image the quality phantom used in some of the Genisys studies. All dimensions are in millimeters. Activity concentration within one of the large cylindrical chambers was approximately 5x that of the rectangular pool. The other large cylindrical chamber was filled with inactive saline. The five rods (varying in diameter between 1 mm and 5 mm) consisted of an activity concentration equal to that of the pool.

This phantom was employed when testing the Genisys for uniformity, contrast-to-noise ratio (CNR), and resolution recovery coefficient. Using a solution of ^{18}F-Fluorodeoxyglucose (^{18}F -FDG), a single cylindrical chamber, all rods, and the rectangular pool were filled such that the cylindrical chamber contained an activity concentration (24.5 μCi/mL) that was approximately 5.3 higher than that of the rods and pool (4.6 μCi/mL). The second cylindrical chamber was filled with inactive saline. The total activity within the entire phantom was approximately 74 μCi at the start of acquisition. Events were acquired over an 8 hour period resulting in a very high-statistic data set (~379 million true events). Histogramming was performed using decay correction, randoms correction, deadtime correction, and detector

normalization. Partial volume correction, scatter correction, and attenuation correction were not available on the Genisys system. A series of static images were then reconstructed using 10, 20, 30, 40, 50, 60, 70, 80, 90, and 100 iterations with the default EM algorithm. A post-reconstruction Gaussian smoothing filter of standard deviation 0.35 mm was applied post-reconstruction. Regions of interest (RIOs) were drawn about each section of the phantom. Note that the ROI corresponding to the rectangular pool is not drawn to contain the entire pool; the pool ROI was drawn to include signal from the pool itself while avoiding contamination from the rods and hot chamber. The smallest diameter rod (1 mm) was not visible in any image. An example of this series is shown in Figure 2.4.
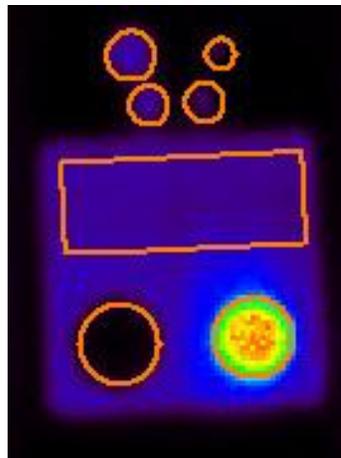
Figure 2.4: Coronal image of phantom with ROIs drawn to segment the various sections. The hot cylindrical chamber is on the lower right. The rectangular pool physically encompasses both cylindrical chambers, but its ROI is drawn to prevent contamination from the rods and hot chamber. Four of the five rods are visible in the upper central portion.

Note that the hot chamber (Figure 2.4) appears to have higher activity concentration within its interior compared to its periphery. The lack of attenuation correction should render this to be the opposite. This appears to be one of the artifacts of the limited angle tomography. In

fact, it is clear that the signal within the hot chamber "spills over" into the relatively cold background pool.

In order to test the effect of low statistics in conjunction with the differing number of iterations, the bootstrap method was employed. In addition to reconstructing with varying number of iterations, the image was also reconstructed using a randomly chosen subsample of the original true events data set. The subsamples included 1%, 10%, 25%, 50%, and 75% of the full data set. Subsampling was performed by reading the listmode file event by event within IDL; prompt and random events were kept or discarded based on the output of a random number generator adjusted to keep a certain percentage of the overall events. Single events were not discarded. The retained events were then written to a new listmode file. This file was subsequently processed using the Genisys histogramming and EM reconstruction protocols. The subsampled set was independently created four times for each percentage. An example of an image created from a subset is shown in Figure 2.5.
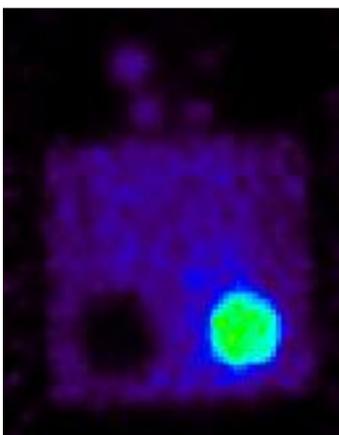
Figure 2.5: Coronal image of phantom using 1% of the data used to create the image shown in Figure 2.4. The same number of iterations (50) and smoothing filter was used as that for Figure 2.4. Four of these subsampled images were created independently for each percentage and each iteration.

Note that the 2 mm diameter rod is barely visible in Figure 2.5 due to the limited number of true events (1% of ~379 million). However, the interior of the hot chamber appears to be more uniform. Spillover from the hot chamber appears to be better contained as well. Based on the examples in Figure 2.4 and Figure 2.5, it appears that a balance between total counts and number of iterations may be struck in order to achieve optimal image quality. A low number of counts does not necessarily imply that more iterations are needed for optimal quality.

2.4   Uniformity Measurements

Using the four images obtained by subsampling at fixed percentage and with a fixed number of iterations, a "mean image" and "standard deviation image" were computed. An example of this is shown in Figure 2.6.
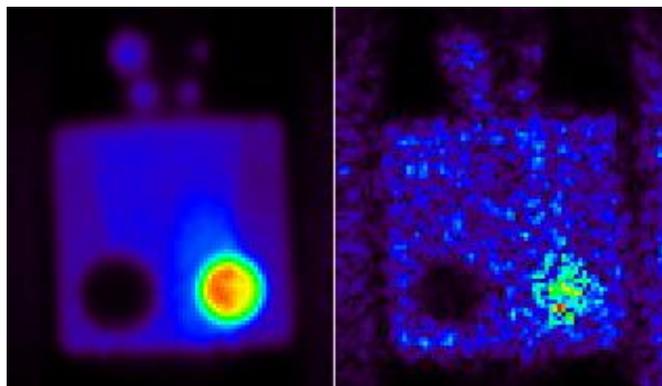
Figure 2.6: Mean image (left) and standard deviation image (right) computed for the phantom using four subset images each created with 1% of the data. One of the four images is shown in Figure 2.5.

After reconstructing these mean images and standard deviation, average values within each region were computed using the ROIs shown in Figure 2.4. The average ROI values were then plotted against number of iterations for each fixed subsample percentage. Computations were not computed when using 100% of the data due to theoretical statistical inconsistencies. Figure 2.7 through Figure 2.10 depict the average ROI values from these images for all subsets plotted against the number of iterations.
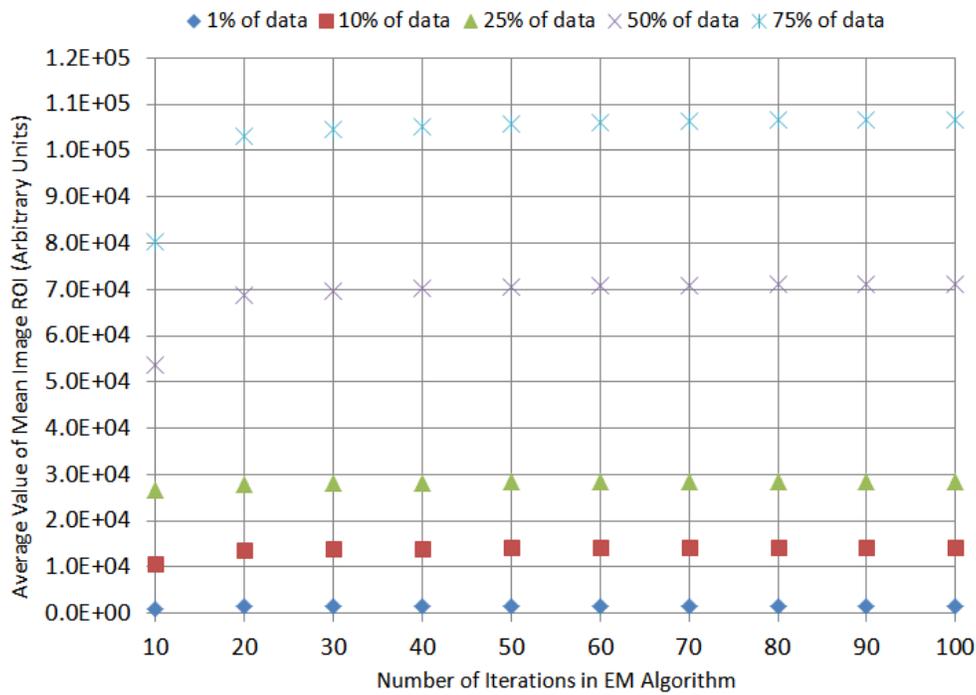
Figure 2.7: Average hot chamber ROI values derived from mean images. Regardless of how much of the original data set was used, there is relatively little change after ~30 iterations.
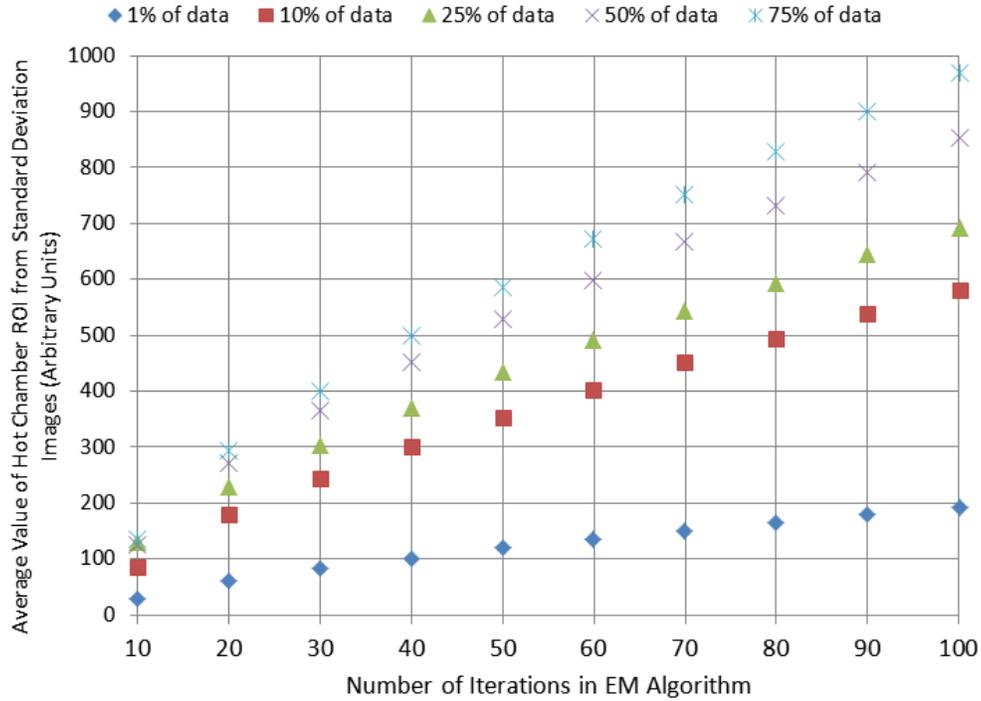
Figure 2.8: Average hot chamber ROI values derived from the standard deviation images. It is clear that these values increase steadily with iterations, especially for the sets with a higher percentage of the original data.
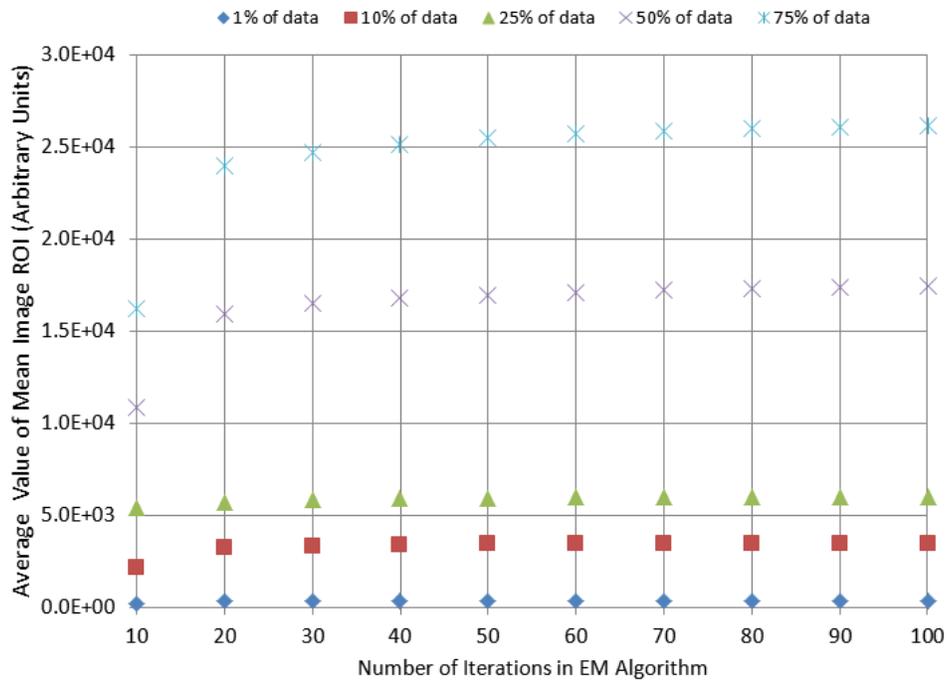
Figure 2.9: Average pool ROI values derived from mean images. Average values are relatively consistent after ~40 iterations regardless of how much data is used.
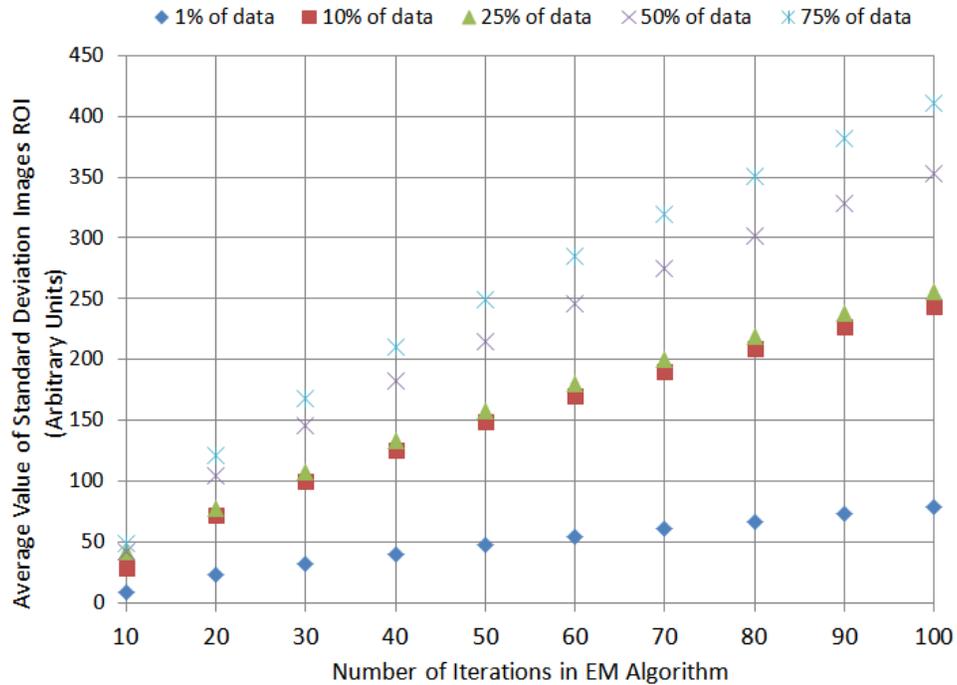
Figure 2.10: Average pool ROI values derived from standard deviation images. . It is clear that these values increase steadily with iterations, especially for the sets with a higher percentage of the original data.

Data from the mean images, as seen in Figure 2.7 and Figure 2.9, indicate that the mean values within each ROI do not significantly change when increasing the number of iterations beyond 30. Data from the standard deviation images, as shown in Figure 2.8 and Figure 2.10, show markedly increased deviation when using a higher number of iterations. This is highly undesirable for ROIs with uniform activity distributions.

2.5   Contrast-to-Noise Ratio

The phantom data set also allowed for a quantitative measure of the contrast-to-noise ratio (CNR). For a given percentage of data and given number of iterations, contrast was defined as

$$Contrast = \frac{Mean\ of\ Hot\ Chamber\ ROI - Mean\ of\ Pool\ ROI}{Mean\ of\ Pool\ ROI} \qquad \text{Equation 2.1}$$

The ROIs used to compute contrast are drawn from the mean images. Noise was defined as

$$Relative\ Noise = \frac{Mean\ of\ Pool\ ROI\ from\ Standard\ Deviation\ Image}{Mean\ of\ Pool\ ROI\ from\ Mean\ Image} \qquad \text{Equation 2.2}$$

Note that noise is defined relative to a mean. Both contrast and noise are dimensionless quantities. Ratios of contrast-to-noise were computed and plotted against number of iterations for each percentage, as shown in Figure 2.11.
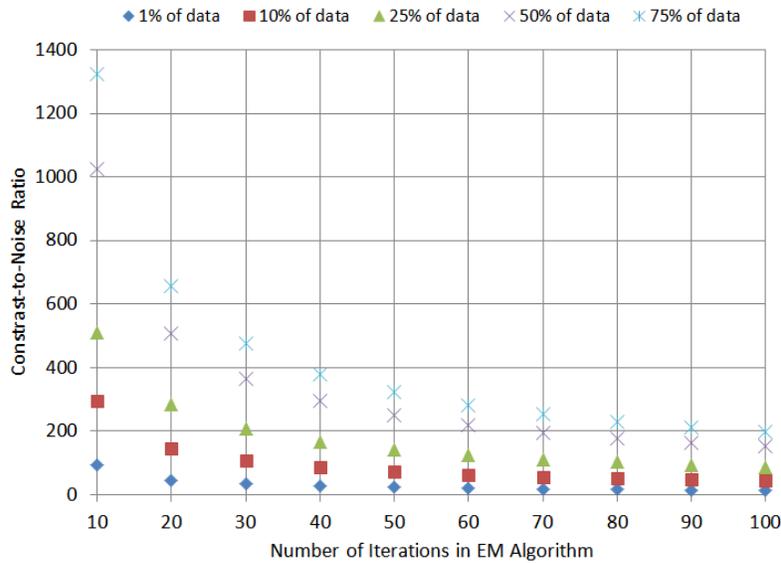
Figure 2.11: Contrast-to-noise ratio measured using ROIs from hot chamber and pool. Contrast itself was found to be slightly lower than expected, but relatively constant regardless of the

number of iterations. The noise, however, increased steadily with iterations. As a result, the CNR decreased rapidly with number of iterations.
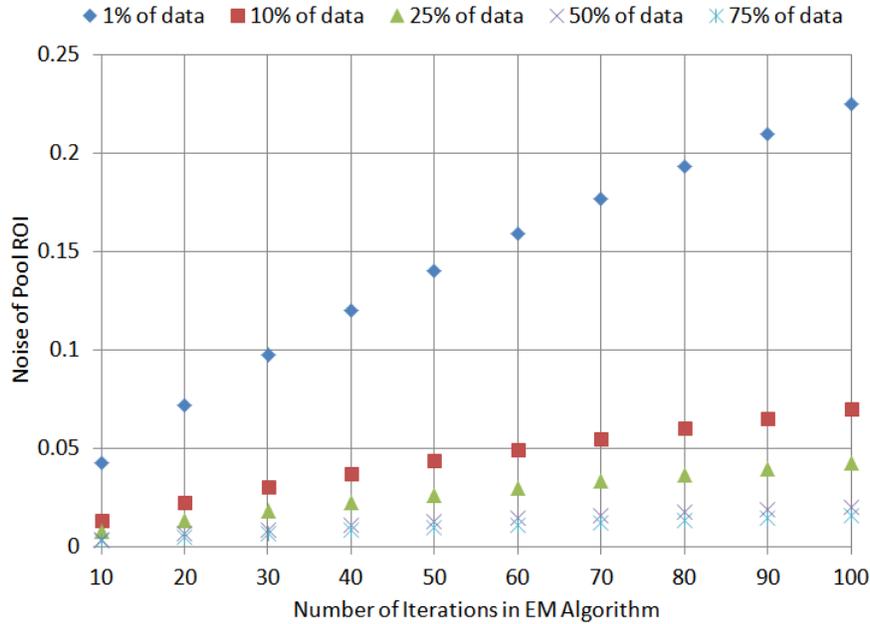
Figure 2.12: Relative noise derived from pool ROI values. The relative noise increases in a linear fashion with the number of iterations, regardless of how much of the original data set was used.

As can be seen Figure 2.11, the CNR generally decreases as the number of iterations is increased. Contrast itself was observed to be somewhat lower than expected based on the relative activity concentrations of the hot chamber and the pool. Furthermore, while contrast itself was observed to remain fairly constant regardless of percentage or number of iterations, relative noise decreased significantly with higher counts as seen in Figure 2.12. Therefore, these CNR results appear to be consistent with qualitative expectations.

2.6   Recovery Coefficients

Spillover between regions with significantly different activity concentrations may affect the accuracy of the voxels values within these regions [6]. The recovery coefficient is a quantitative measure of how well true voxel values are retained when neighboring voxels are of a different activity concentration. Recovery coefficients for all visible rods were computed. Each rod was physically separated from other regions. For each rod acquired at a given subsample percentage and iteration, a centroid within its interior was computed based on the mean images. All voxels within a neighborhood of 2 voxels about this centroid were averaged. The ratio of this value to that of the mean ROI value of the entire rod, as shown in Figure 2.4, is then defined to be the recovery coefficient:

$$Recovery\ Coefficient = \frac{Mean\ of\ Voxels\ Near\ Centroid\ of\ Rod}{Mean\ Value\ of\ Rod\ ROI} \qquad \text{Equation 2.3}$$

Since the 1 mm diameter rod was not visible in any of the images, it was not included in this analysis. The recovery coefficients of the remaining rods for each of the percentages are shown in Figure 2.13 through Figure 2.18.
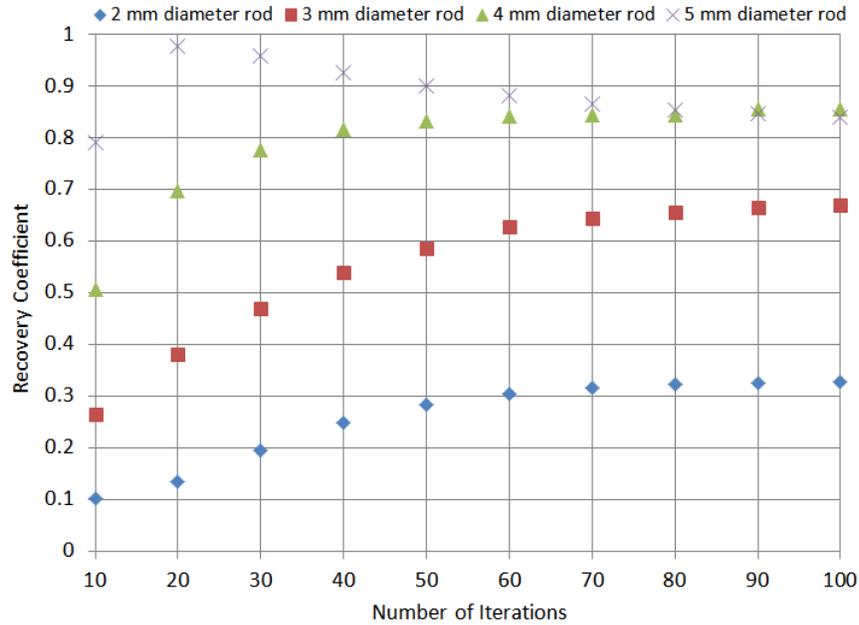
Figure 2.13: Recovery coefficients of rods using 1% of data. The 1 mm diameter rod was not visible in the final image. All recovery coefficients are less than 1, indicating that the spillover effect is dominant regardless of the number of iterations used.
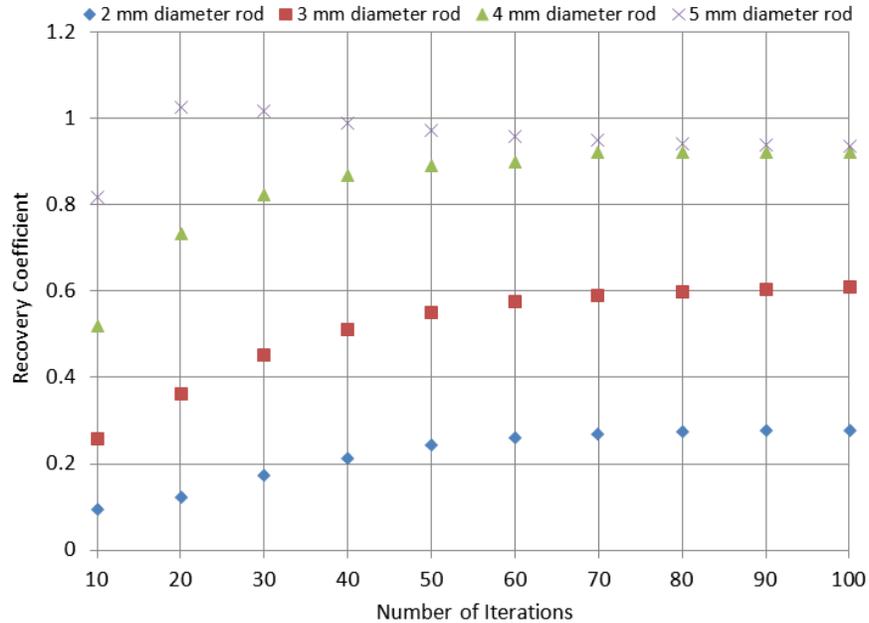
Figure 2.14: Recovery coefficients of rods using 10% of data. The 1 mm diameter rod was not visible in the final image. Most coefficients are less than 1. However, coefficients greater than 1 are possible due to spillover of activity from the periphery of the rod into the centroid.
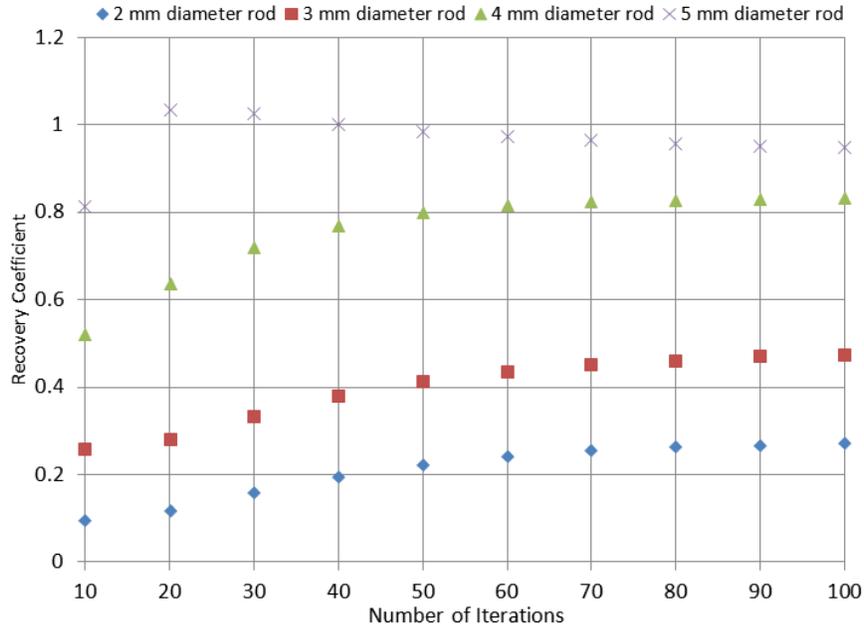
Figure 2.15: Recovery coefficients of rods using 25% of data. The 1 mm diameter rod was not visible in the final image. Most coefficients are less than 1. However, coefficients greater than 1 are possible due to spill over of activity from the periphery of the rod into the centroid.
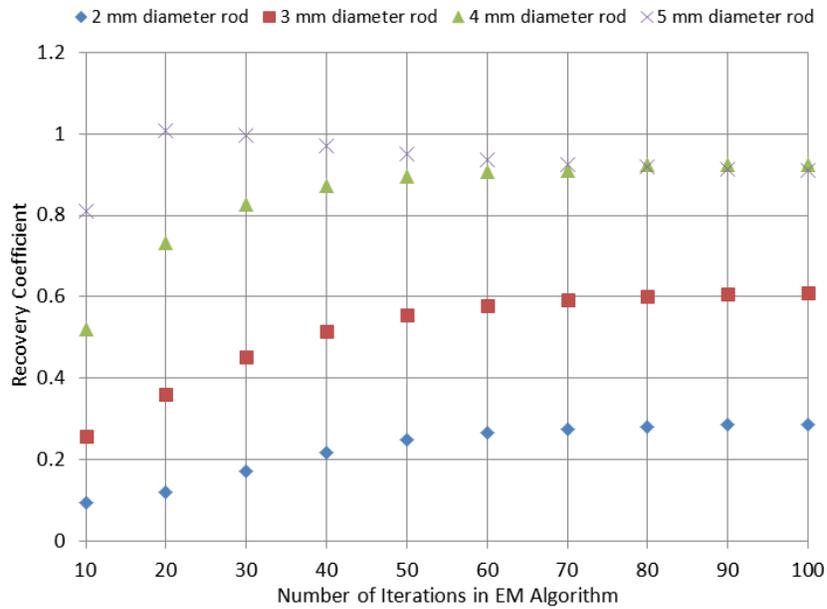
Figure 2.16: Recovery coefficients of rods using 50% of data. The 1 mm diameter rod was not visible in the final image. Most coefficients are less than 1. However, coefficients greater than 1 are possible due to spillover of activity from the periphery of the rod into the centroid.
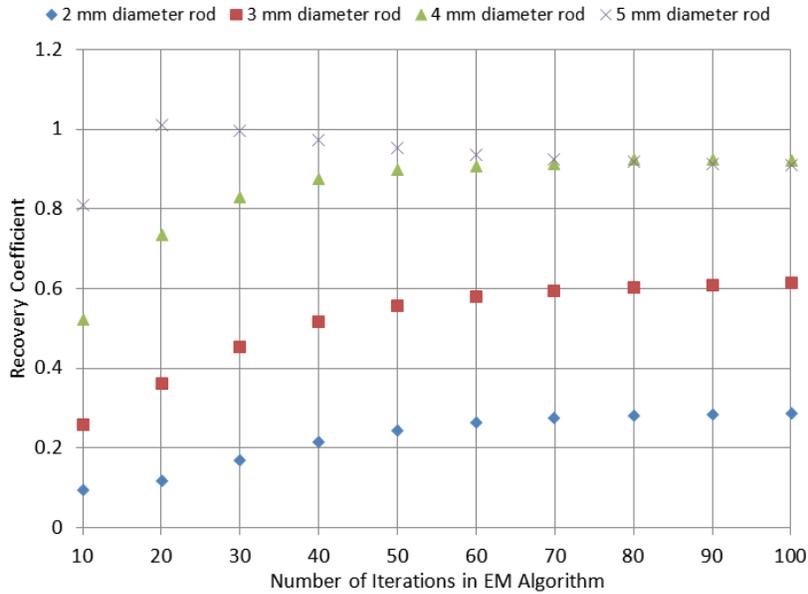
Figure 2.17: Recovery coefficient of rods using 75% of data. The 1 mm diameter rod was not visible in the final image. Most coefficients are less than 1. However, coefficients greater than 1 are possible due to spillover of activity from the periphery of the rod into the centroid.
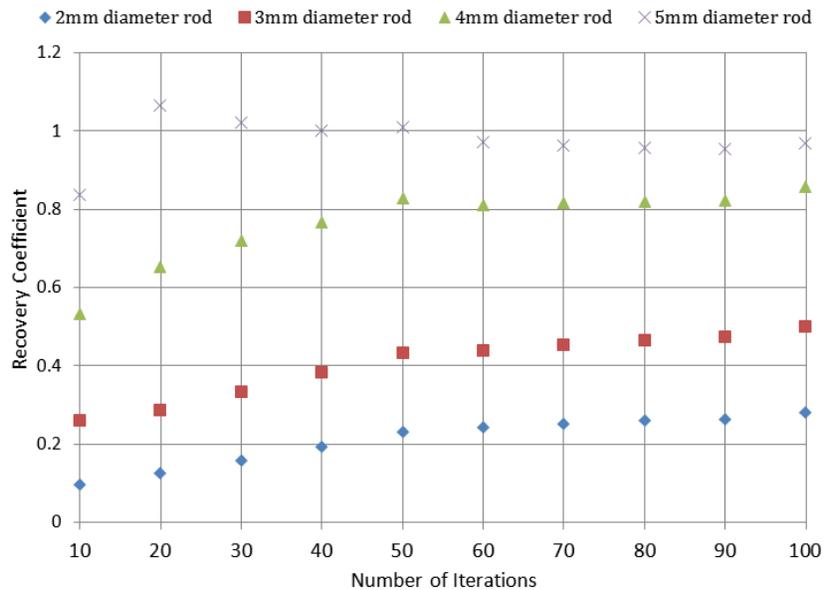
Figure 2.18: Recovery coefficients of rods using 100% of data. The 1 mm diameter rod was not visible in the final image. Most coefficients are less than 1. However, coefficients greater than 1 are possible due to spillover of activity from the periphery of the rod into the centroid. There appears to be slightly more fluctuation in the coefficients when using this complete data set compared to the randomly selected subsets.

21

The recovery coefficients tend to increase slightly with increasing number of iterations. The 5 mm diameter rod appears to be an exception to this, but remains relatively unchanged after reaching an initial peak at ~20 iterations. In addition, a recovery coefficient of greater than 1.0 for this rod has been measured in the low statistic data. This overestimation of voxel values may be due to spillover of activity from the periphery of the rod into its interior. The spillover of activity outside of the rod ROI appears to be dominant for rods of diameter less than 5 mm.

2.7   Deadtime Measurement

As is common with most nuclear medical systems, the Genisys is prone to loss of counts during periods of excessively high count rates known as deadtime. Deadtime is typically due to signal "saturation" within the photomultiplier tube amplifiers; in short, counts may not be registered due to the system being temporarily "paralyzed" during deadtime. Deadtime is measured using the singles count rate over the entire FOV. Therefore, only a global estimate of deadtime is typically measured [7].

The total activity at which this global deadtime becomes intolerable is of interest in order to avoid loss of useful event acquisition. A continuous scan of relatively high initial activity is performed such that the duration of the scan allows for the nearly complete decay of all activity. For measurements performed on the specifically on the Genisys, a uniform cylinder ($\emptyset$ = 25.5 mm, length of 76.0 mm, fluid volume of 21 mL) was filled with approximately 102 μCi of ^{18}F – FDG activity. A 10-hour scan was performed and histogrammed into 20 30-minute frames using the default 3D protocol.

Note that the deadtime correction is applied to counts within the sinogram, as opposed to voxel values within the image. As a result, image reconstruction was not necessary. Two data sets were created. The first set of 20 sinograms included normalization and randoms correction, but lacked decay correction (scatter correction and attenuation correction were not available); the second set of 20 sinograms included the deadtime correction built into the Genisys histogramming protocols. As a result, elements within the sinogram are interpreted as the total number of events (singles, randoms, and true coincidence detection of annihilation photons) recorded along a given LOR.

At low activity levels, the induced deadtime is considered to be negligible; in other words, the observed count rate can be considered to be a very good approximation to that of the true count rate. The true count, per frame, at higher activity levels can then be extrapolated from this low-activity regime using at least two measured points and the initial activity. Due to the exponential nature of the decay, the extrapolation is nonlinear. However, the data can be put into a linear form when the counts are plotted semi-logarithmically against frame time (as opposed to activity level). Once the extrapolation is performed, the true counts may be computed at any activity level. For the results shown here, only the last two points were used for the extrapolation. The quantity of interest is the activity level at which the observed number of counts is approximately 20% lower than that of the true number of counts. The true counts themselves are estimated from the extrapolated counts.

The semi-logarithmic plot is shown in Figure 2.19, and illustrates the linear extrapolation from the low-activity regime (i.e. frame time at 9.0 hours and 9.5 hours) towards the high-activity regime (i.e. earlier frame times).
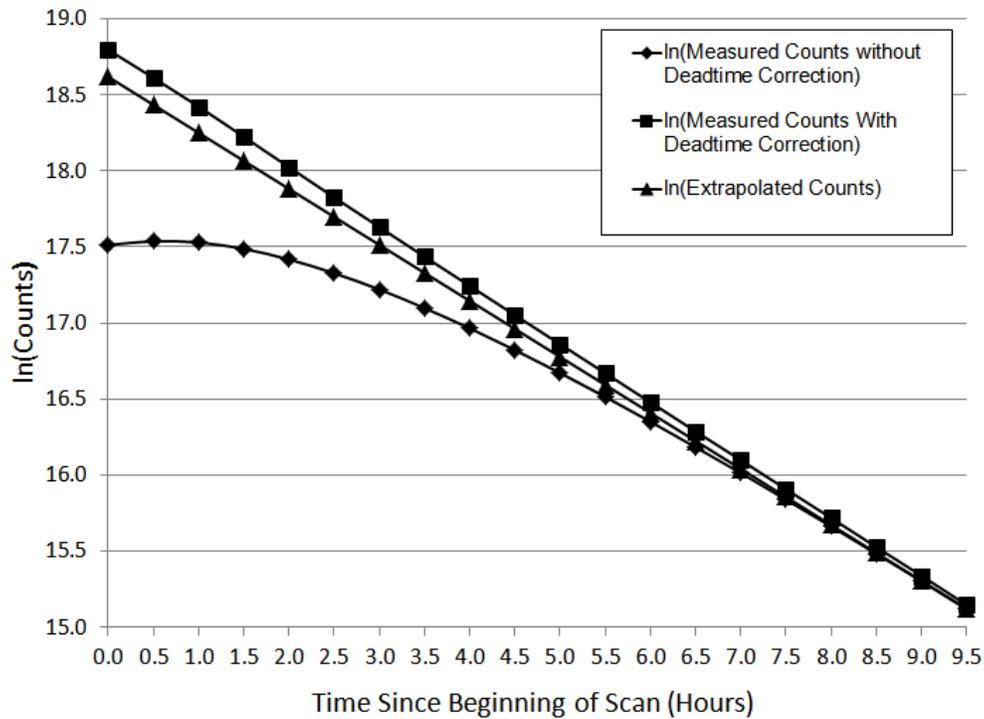
Figure 2.19: Semi-logarithmic plot indicating the measured counts without deadtime correction and the extrapolated true counts. The measured counts with deadtime correction are shown only for visual comparison; it was not used within the analysis.

Prior to the 0.5-hour mark, the system is unable to accommodate all events; after the 0.5-hour mark, the number of events gradually becomes more manageable. The extrapolated counts curve represents the number of counts measured, per frame, if the system did not suffer from any deadtime or any other nonlinearity.
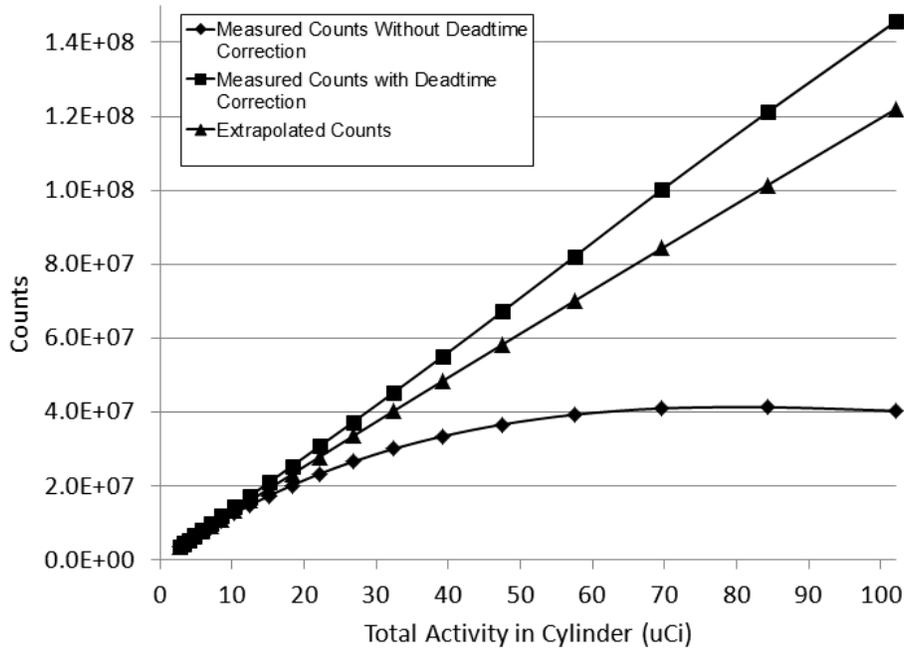
Figure 2.20: Counts acquired as a function of activity within the Genisys system. The activity indicated is based on the half-life of ^{18}F (109.8 minutes) and the initial activity (102 µCi) as measured by a dose calibrator prior to the scan. The true counts were extrapolated from the region of lowest activity where deadtime is the least intrusive. The measured counts with deadtime correction are shown only for visual comparison; it was not used within the analysis.

The counts-vs-activity plot corresponding to Figure 2.19 is shown in Figure 2.20. Figure 2.20 presents the same information as that in Figure 2.19, but allows determination of the activity at which the measured counts fall to approximately 20% of the true counts. As seen in Figure 2.20, this occurs at an activity level of approximately 22 µCi. This discrepancy will only increase at activity levels greater than this, and vice-versa. Note that the high activity region within Figure 2.20 corresponds to the low frame number region in Figure 2.19 due to the fact that the scan was performed with a relatively high activity initially, and then allowed to naturally decay as time progressed. In addition, application of the deadtime correction within the Genisys appears to overestimate the number of true counts at nearly all activity levels. Most scans performed on the

Genisys, however, are performed at relatively low activity levels (between ~20 μCi and ~30 μCi). As a result, this discrepancy is minimized.

Figure 2.20 also provides a rough estimate for the width of the deadtime window (in units of time). The activity level of approximately 91 μCi represents a "critical point" for the system. For activity levels lower than this, the timing between consecutive detection events are long enough for the system to manage; although counts may be lost at this activity level, the measured count rate is proportional to the true count rate. As a result, the average time between detected events at 91 μCi represents the shortest window for which the system may adequately accommodate detection events. This can be found simply by converting the activity into units of counts-per-second (cps) and then taking the reciprocal. The resulting value is 29.5 μs. In other words, activity levels at which consecutive detection events occur, on average, within 29.5 μs of each other will induce paralyzable deadtime within the system. Detection events which take place more than 29.5 μs apart will induce deadtime, but not paralyze the system. Again, this estimate of the deadtime window should only be considered as a rough approximation. A more accurate calculation would require the instantaneous count rate over time in order to find the "critical point". In addition, only the average time between events is considered; as can be deduced from Poisson statistics and error propagation, lower activity levels will inherently have a higher average timing window (and higher variance of this window) between detection events compared to that of higher activity [7].

2.8   Summary

A number of performance characteristics for the Genisys have been measured. Spatial resolution was measured at a moderate ~1.5 mm while CNR and resolution recovery proved adequate. Deadtime was found to be somewhat restrictive; even a moderate activity of 22 μCi induced about a 20% loss of counts. Even though this may be compensated for using deadtime correction, the correction itself is a somewhat crude global correction. As result, regions with a relatively high activity will be under-corrected, and vice-versa.

Along with the image quality phantom, measurements for SNR and uniformity were measured using the bootstrap method. Note that this method acquires subsamples from a given "global" distribution in order to approximate a data set from a limited number of samples. For the work presented here, four such sets were created. However, it is unclear if this is an adequate number of samples; even global distributions generally require more samples to find a representative data set compared to uneven distributions (as seen in information theory). This issue may be avoided by applying analytical statistical methods to the global distribution in order to obtain the most representative data set without the need for subsampling. Future work may also investigate a direct measure of parallax and intercrystal scattering by measuring line widths within the sinogram. The flat geometrical configuration of the system makes this investigation much more convenient than that performed on a typical cylindrical system in which the curvature of the crystal elements may interfere with an accurate measurement.

3   Factor Analysis on the Genisys System Images

3.1   Introduction

Due to high cost of traditional preclinical positron emission tomography (PET) systems, recent advances in low-cost, simplified systems have been made. One such system, Genisys (Sofie Biosciences, Culver City, CA 90230), utilizes two planar bismuth germinate (BGO) crystal arrays [1, 2]. Each 5-mm thick array includes 20 x 44 crystal elements with dimensions of 2 mm x 2 mm. Each BGO array is coupled to two position-sensitive photomultiplier tubes (PSPMT) (Hamamatsu Photonics, Bridgewater, NJ). The upper and lower detector heads are separated by 5 cm and measure 5 cm x 10 cm each as seen in Figure 2.1.

Due to the lack of data from certain lines-of-response (LORs), the scanner suffers from limited tomography. This limited tomography, in turn, results in relatively poor spatial resolution particularly for the transaxial and sagittal planes. This poor spatial resolution then results in the spillover effect in which activity from one region intrudes on activity within another region, thereby preventing accurate quantification of activity within a given region. The spillover in the coronal plane is minimized due to the geometrical configuration, but still prevalent.

Although cost-effective and suitable for distant-organ imaging, poor resolution and delineation of neighboring organs may become unreliable as a result of the limited angle tomography. This is particularly true for cases in which structures are aligned vertically (AP direction). This is illustrated in Figure 3.1, where two sources of approximately the same activity are placed next to each other in the horizontal plane (left) and in the vertical plane (right).
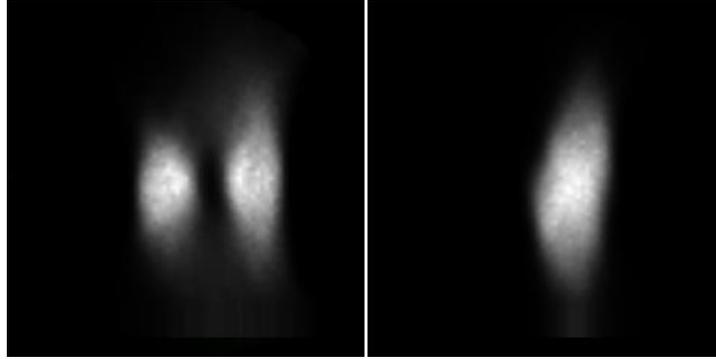
Figure 3.1: Transaxial slice of horizontally (left) and vertically (right) aligned chambers of approximately equal activity. Due to the limited angle tomography, clear separation is only possible when the sources are horizontal relative to each other.

As can be seen in Figure 3.1, when the sources are adjacent to each other in the horizontal plane the sources can be easily distinguished, although distorted due to limited angle tomography reconstruction. When the sources are placed on top of each other it is virtually impossible to distinguish the two sources.

Factor analysis (FA) has been proposed as a method to mitigate this. Previous work has shown FA to be a suitable statistical method of organ delineation in dynamic imaging studies utilizing the different tracer kinetics of different tissue types [8, 9, 10, 11]. The work presented here attempts to address the issue of organ delineation on the Genisys with the application of FA on the reconstructed dynamic images.

When interpreting the results of FA, a few key points should be kept in mind. Ideally, a set of a few factor curves should retain all the information in a given set of measured TACs. A quantitative measure of how much information is retained is given by the percentage of total variance explained by the factor curves. This is loosely analogous to the Modulation Transfer Function (MTF) used when assessing image quality. The variance, as used here, is in reference to total temporal variance of all measured TACs. It is not in reference to the variance of the image

29

in the spatial domain. In addition, the percentage of explained variance associated with each factor curve is not a measure of how well FA separates structures. Factor curves are often ordered such that the factor which retains the largest percentage of the original data is first. Successive factor curves will explain a successively smaller percentage. When attempting to calculate the overall percentage of variance explained by all factor curves, it is not entirely accurate to add the individual percentages; the nonzero covariance between factor curves may increase or decrease substantially depending on the data and number of factor curves chosen to represent them. From each factor curve, a corresponding factor image may be computed. These factor images indicate which voxels have a high degree of resemblance, over time, to the corresponding factor curve. The factor images are not dynamic images; they simply indicate the linear combinations that are used to construct the measured TACs from the factor curves. Factor images should be thought of as weighting maps.

3.2   Methods

3.2.1 Calculation of Factors

The number of factors used to represent a given data set is primarily limited by computational requirements. A large number of frames will typical preclude the use of many factors. For our purposes here, only two or three factors were used for a given study. The method of determining the factors is based on Di Paola [9]. A least-squares criterion is used as the cost function to be minimized. Note that, unlike other methods such as Principal Component Analysis (Chapter 4), all desired quantities must be calculated simultaneously. The elements of all factor

curves should be interpreted as a set of variables that minimize the cost function. The least-squares criterion allows for the problem to be reduced to a set of linear equations. These equations are then iteratively computed using the Nelder-Mead simplex method [12, 13, 14]. A non-negativity constraint is applied to each factor in order to obtain physically reasonable results.

As with most iterative methods, convergence of a solution is paramount. In addition, the choice of initial factors may have an impact on the number of iterations needed for convergence. For these purposes, the initial factors are computed based on Principal Component Analysis. Not only do the principle components serves as initial factors, but the percentage of data retained by using these factors, regardless of the number of iterations computed, is also calculated. This relationship is further exploited in Chapter 4.

3.2.2 Dual-Isotope Experiment

The ability of FA to spatially separate structures relies heavily on the relative covariances between TACs. These TACs, in turn, rely not only on the physiological tracer kinetics, but also on the noise inherent in the system. In order to quantitatively assess the use of FA in the Genisys, a dual-isotope scan was performed. The explained variance of the resulting two factor curves can then be considered as a metric for how much information is retained.

The dual-isotope scan provides an ideal testing ground for FA in a few different ways. First, the activity concentration at a given time can be calculated based on the isotopes' half-lives and initial concentration values. Using these computed curves, the covariance between TACs can also be estimated. Secondly, the use of only two linearly independent sources is amongst the

simplex problems for which FA can be applied. Ideally, only two factor curves should account for 100% of the temporal variance in the dual-isotope image. Noise, however, may severely reduce this percentage. Note that the effectiveness of FA is not dependent on spatial relations between sources. However, the use of two spatially separate sources also allows visual comparison between the factor images and reconstructed images. Finally, the use of spatially separated sources with very limited volume limits any spillover between sources. This also aids in visual comparison between reconstructed images and factor images.

For the purposes of this experiment, two vials were constructed in order to hold tracer of high concentration, but negligible volume. These "point sources" consisted of one vial containing ^{18}F ($T_{1/2}$: 109.8 minutes), and another vial containing ^{11}C ($T_{1/2}$: 20 minutes). At the beginning of the 30-minute acquisition, the ^{18}F point source contained approximately 22 µCi of activity while the ^{11}C point source contained approximately 29 µCi of activity. Dynamic reconstruction included deadtime correction, randoms correction, and normalization, but did not include decay correction or attenuation correction. Reconstruction of the dynamic images was performed using 30 frames. All frames were reconstructed using the Expectation Maximization (EM) Algorithm with 100 iterations and a post-reconstruction Gaussian smoothing filter (σ = 0.35 mm). Adjacent frames were spaced 1 minute apart, but had a variety frame durations including: 1-minute; 30-second frames; 15-second; 10-second; and 1-second. These frame durations were designed to reduce the signal-to-noise ratio (SNR), yet maintain the same overall kinetics. An example of a given frame with ROIs is shown in Figure 3.2.
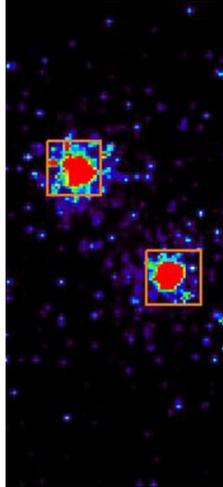
Figure 3.2: Example of coronal slice from the dual-isotope experiment. The image shown corresponds to the first frame of the 15-second data set. The color mapping has been chosen to emphasize noise surrounding the point sources. The ^{11}C source is on the upper-left; the ^{18}F source is on the lower-right.

The deterioration of the exponential decay curves, due to this reduction in SNR, is illustrated in Figure 3.3 and Figure 3.4. Noise-free exponential decay curves, with the same initial activities as those shown in Figure 3.3 and Figure 3.4, can be computed. Based on this simple model, the inherent variance of the ^{11}C curve is approximately 1.4 times that of the ^{18}F.
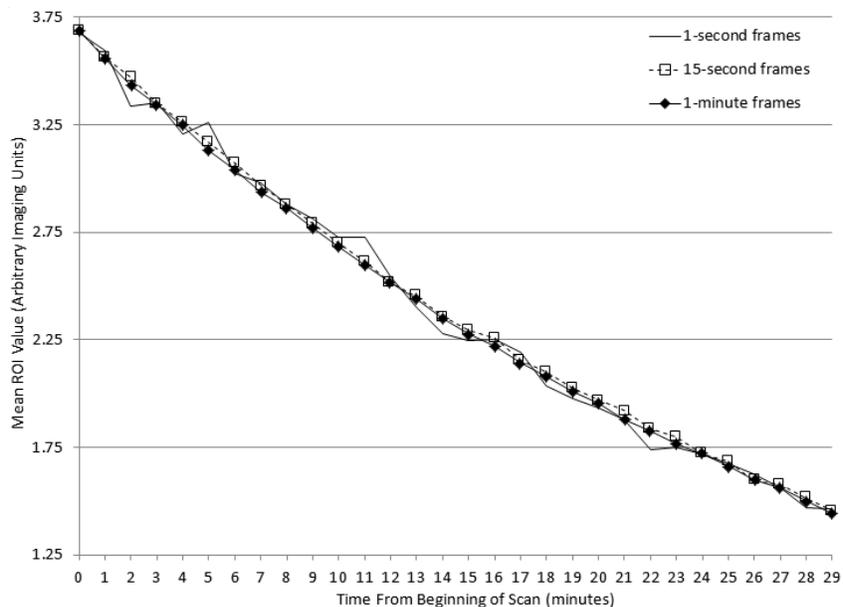
Figure 3.3: ^{11}C decay curves measured from the raw dynamic images pertaining to the ^{11}C ROI shown in Figure 3.2. The 1-minute data set makes full use of all acquired data and is, therefore, the least noisy.
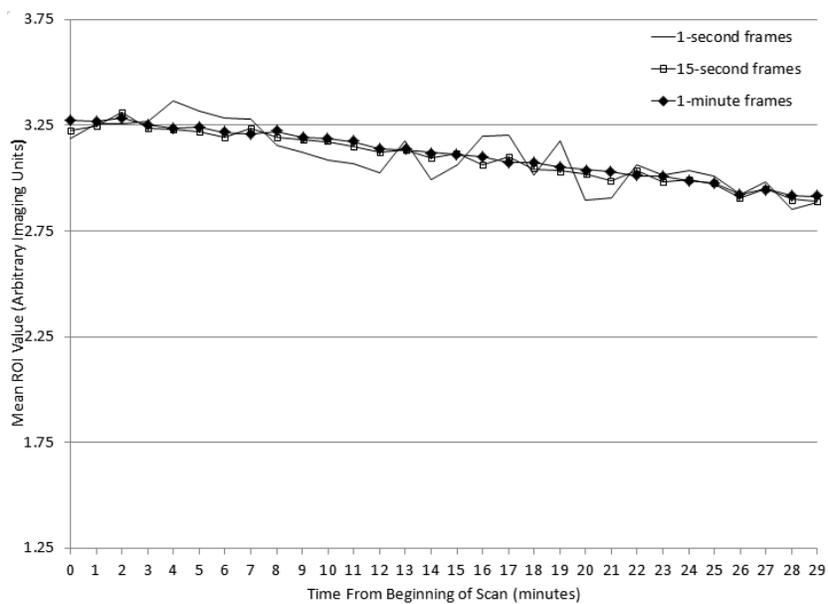
Figure 3.4: ^{18}F decay curves measured from the raw dynamic images pertaining to the ^{18}F ROI shown in Figure 3.2. The 1-minute data set makes full use of all acquired data and is, therefore, the least noisy. The scaling is set to match that of Figure 3.3 for visual comparative purposes.

For each dynamic set, FA was applied to the raw Genisys images. Each raw image contains 80 x 80 x 176 voxels (0.55 mm x 0.55 mm x 0.55 mm). Amongst all these voxels, only 1330 TACs can be used in FA due to computational limitations. Therefore, the 665 TACs with the largest variance were chosen from each of the two ROIs. This selection process not only results in a manageable subset, but also compensates for differences in the number TACs corresponding to each isotope. By choosing equal numbers of TACs from both regions, results from FA will be based purely on the variances of the TACs and not on the number of TACs acquired from each region. Since FA produces factor curves that resemble TACs of the largest variances, this set of 1330 TACs would prove to be the most influential even if the entire dynamic image were to be processed. Subsets of these 1330 TACs are shown in Figure 3.5 and Figure 3.6.
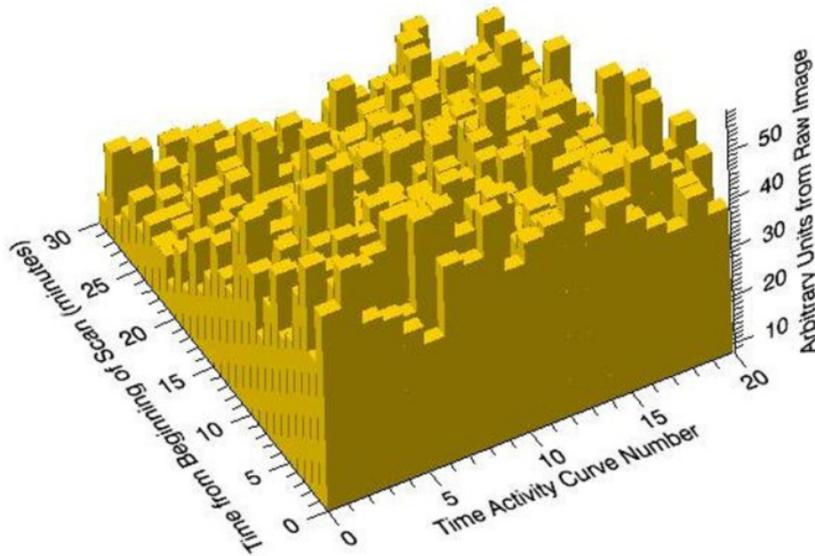
Figure 3.5: Example of voxel-based TACs from dynamic image. TACs derived from the ^{11}C ROI shown in Figure 3.2. The TACs exhibit overall temporal decay, but also exhibit signs of random noise. For visual clarity, only 20 of the 665 TACs selected from the ^{11}C ROI are shown here.
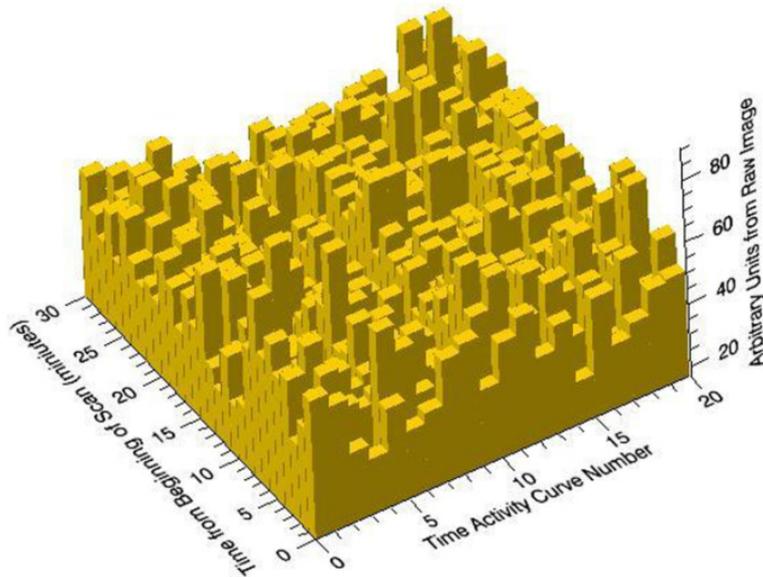
Figure 3.6: Example of voxel-based TACs from dynamic image. TACs derived from the ^{18}F ROI shown in Figure 3.2. The TACs exhibit signs of random noise, but not natural decay due to the relatively long half-life of [18]F (109.8 minutes). The natural decay of ^{11}C shown in Figure 3.5 is more apparent. For visual clarity, only 20 of the 665 TACs selected from the ^{18}F ROI are shown here.

Once the set of 1330 TACs were selected, FA was performed using the Interactive Data Language (IDL) version 7.0. Only two factors were calculated for each dynamic set. The total variance explained by both factors was then calculated and used as a quantitative metric to assess the effectiveness of FA.

3.2.3 Factor Analysis Applied to a Synthetic Data Set

In order to assess the use of FA on the Genisys, a second, semi-quantitative, study was performed. The study involved the use of synthetically created dynamic images. The sets of

dynamic images produced differed in terms of noise and level of overlap between structures. These dynamic sets were individually processed using FA in order to assess the "recovery" of the TACs which were used in the synthesizing of the data. This method consisted of computing two TACs resembling those which may be found during a typical ROI analysis in mice studies. The area under the curve (AUC) of each TAC was normalized to unity. Subsequently, two corresponding static images of high-activity, "noise-free", 3 mL cylindrical chambers (50 mm length and 9 mm diameter) were acquired on the Genisys system. These static images were acquired such that both sinograms contained approximately the same number of total counts. The chambers, both filled with approximately 40 µCi of ^{18}F-FDG, were placed adjacent to one another along the AP direction.

The sinograms corresponding to each image were scaled according to its corresponding synthetic TAC. The resulting sinograms were then summed and reconstructed to form one dynamic data set in which both chambers are present, but have varying temporal behavior. The images were reconstructed using 30 frames with a total scan duration of approximately 50 minutes. The EM reconstruction algorithm with 100 iterations and a post-reconstruction smoothing filter ($\sigma = 0.35$ mm) was used for the reconstruction of all images in the study.

Additional data sets were produced by the same method, but by pruning the original sinograms in order to simulate moderate-statistic image sets. In order to perform the pruning, only a fixed fraction of events from the original listmode file were randomly selected for inclusion in a new listmode file. The number of events which were included within each frame was proportional to the height of the synthetic TAC being used. This sampling method ensured that, while fewer events were captured in the new listmode file, the overall shape of the TAC was preserved. Two new listmode files were created using this method; each one corresponded

to the two synthetic TACs. Sinograms from each of the listmode files were produced and then summed together. This summed sinogram was then reconstructed using EM to produce the final dynamic data set. Note that this method was used to simulate statistical noise present in low activity regions, as opposed to "white-zero-mean" noise. The total number of counts within the moderate statistic data set was only 20% of that within the original high activity data set.

In addition, the degree to which chambers spatially overlapped was also varied, thereby simulating different levels of spillover between the chambers. Each synthetic dynamic data set was then processed with thresholding and curve-smoothing prior to FA. Curve-smoothing was performed by using a 1-2-1 averaging filter for each data-derived TAC. The thresholding and curve-smoothing were performed in order to reduce noise and simplify the data into a more manageable form for the subsequent PCA. Thresholding was applied by retaining values which were within 5% of the maximum value within each data-derived TAC. The overall method is illustrated in Figure 3.7.
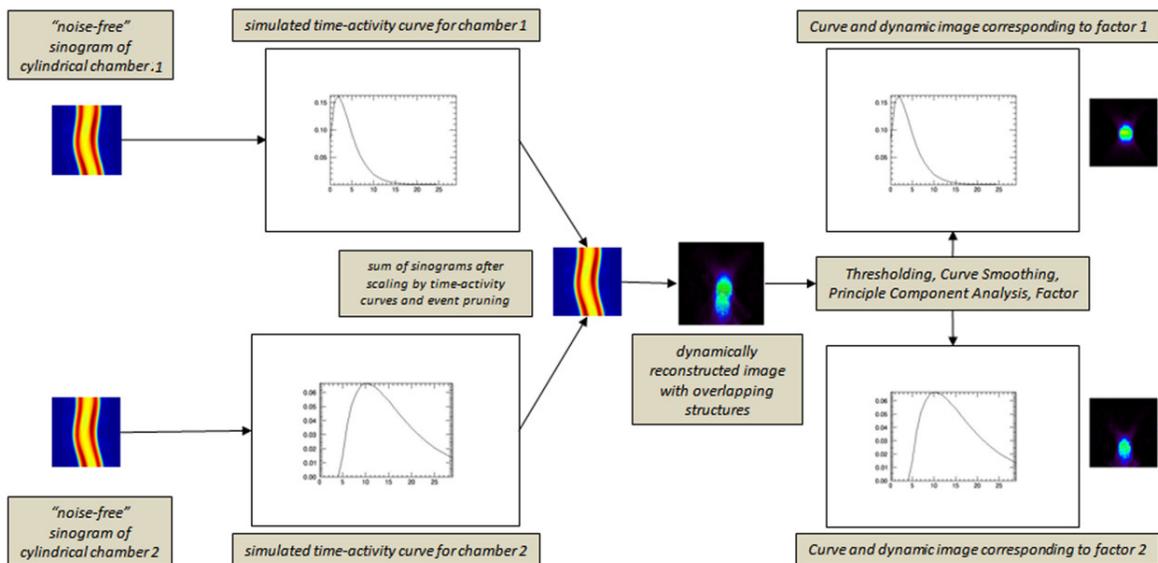
Figure 3.7: Diagram illustrating the overall method used in the production of the synthetic images. Two different synthetic TACs are used to scale the sinograms representing each of the

two chambers. Recovery of these synthetic TACs is attempted using FA. The experiment was repeated twice using images of varying degrees of overlap.

3.2.4 Factor Analysis Applied to a Mouse Study

As an example of the practical application of FA, a mouse study was conducted on the Genisys. The mouse used for this study was treated in accordance with the Animal Research Committee at the University of California – Los Angeles. The mouse, bearing a subcutaneous murine melanoma B16 tumor, was injected with 86 µCi of ^{18}F-FDG. Data was acquired over a 60-minute period. Dynamic reconstruction, employing the EM Algorithm with 100 iterations with post-reconstruction smoothing filter ($\sigma = 0.35$ mm), was performed using 44 frames. Frame lengths were adjusted to properly sample curve kinetics which typically change more rapidly during the early portion of the scan. Factor analysis was then applied to this dynamic set after applying significant thresholding and voxel-averaging in order to reduce noise and render the data more manageable. Note that, for this particular image, activity accumulated within the bladder was grossly disproportionate to the rest of the field-of-view. Consequently, the computed factor curves would be largely biased towards bladder activity. As a result, the bladder, and all inferior regions, were eliminated before Factor Analysis was used. Two factor curves were used in the analysis, resulting in two factor images, and a third "residual" image that is effectively discarded. A summed coronal slice of the mouse image is shown Figure 3.8.
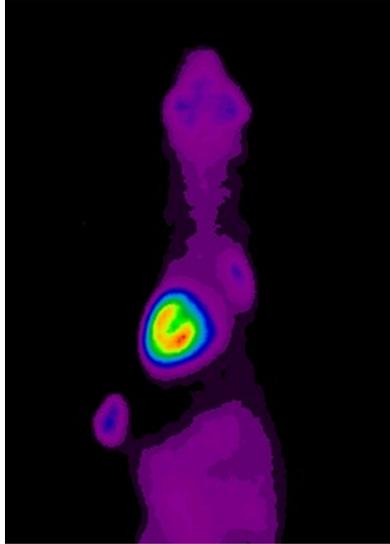
Figure 3.8: Summed-coronal slice of tumor-bearing mouse. The arrow points to subcutaneously injected tumor. The bladder, and regions inferior to the bladder, were removed in order to prevent biasing within the FA results.

3.3   Results

3.3.1 Dual-Isotope Experiment

The explained variances of the dual-isotope experiment are shown in Table 4.

Frame Duration (seconds)	Total Variance of Data Explained by Both Factors
1	7.5%
10	18.0%
15	22.4%
30	32.3%
60	45.4%

Table 4. Percentage of explained variance resulting from Factor Analysis on the dual-isotope dynamic scans.

These results are a quantitative indication of how well two computed factor curves could capture the measured decay curves from two different isotopes. In the ideal, noise-free, case, only two factor curves would capture 100% of the decay from two different isotopes. However, in reality, the noise properties of the system will prevent this. Therefore, these results can be interpreted as how well factor analysis performs in the presence of noise within the system.

3.3.2 Factor Analysis Applied to a Synthetic Data Set

The factor curves derived from the second, semi-quantitative study are shown in Figure 3.9 through Figure 3.12. The factor curves derived from the FA are plotted over the synthetic curves used to scale the sinograms from the images of the two chambers. Note that the calculation of these factor curves did not include any prior knowledge about the synthetic curves. Furthermore, FA was applied to the entire dynamic image which consisted of both overlapping chambers. No analysis was conducted on the individual chamber images or sinograms.
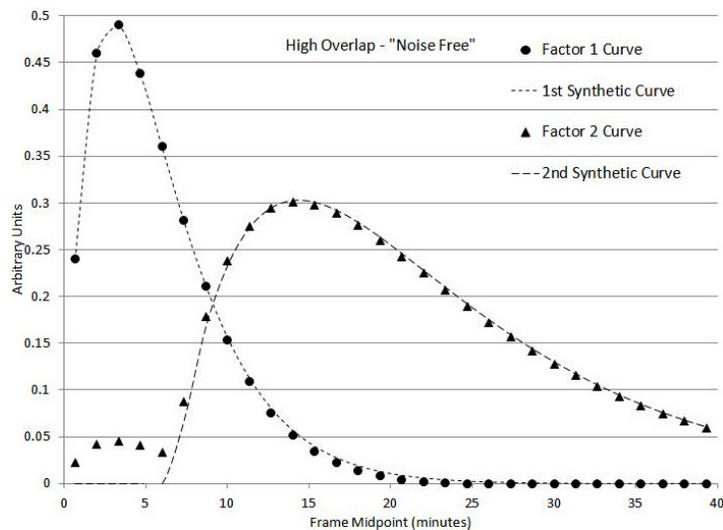
Figure 3.9: Results for the "high overlap" and "noise free" data sets from the semi-quantitative study shown in Figure 3.7. All curves have been normalized to unity using the Euclidean norm.
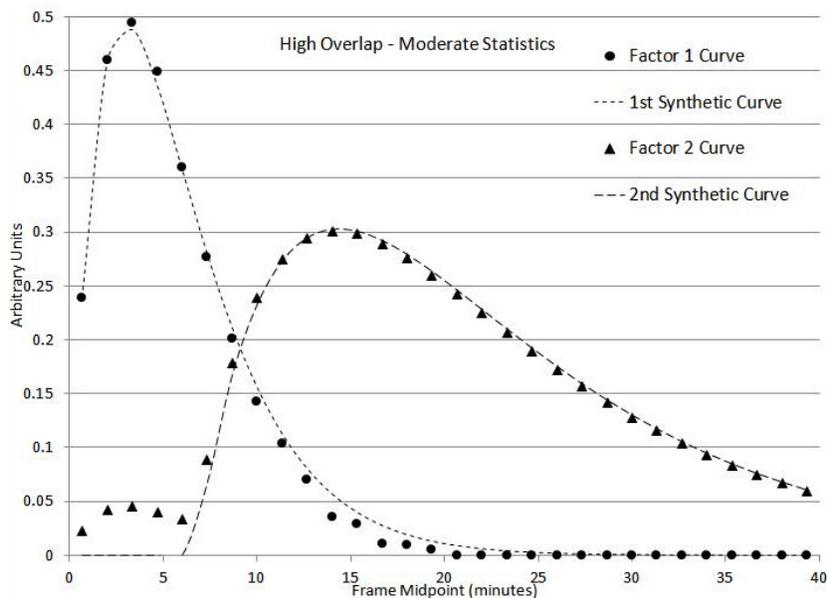
Figure 3.10: Results for the "high overlap" and "moderate statistics" data sets from the semi-quantitative study shown in Figure 3.7. All curves have been normalized to unity using the Euclidean norm.
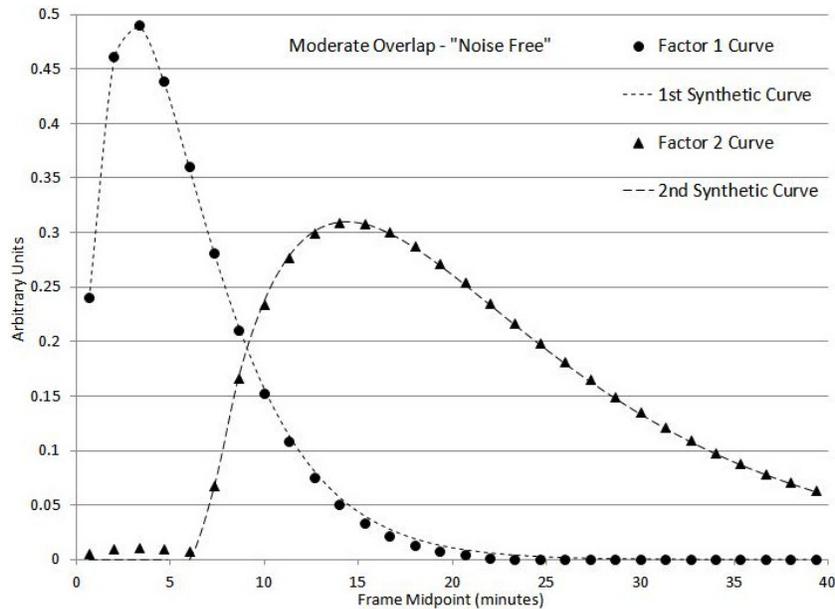
Figure 3.11: Results for the "moderate overlap" and "noise free" data sets from the semi-quantitative study shown in Figure 3.7. All curves have been normalized to unity using the Euclidean norm.
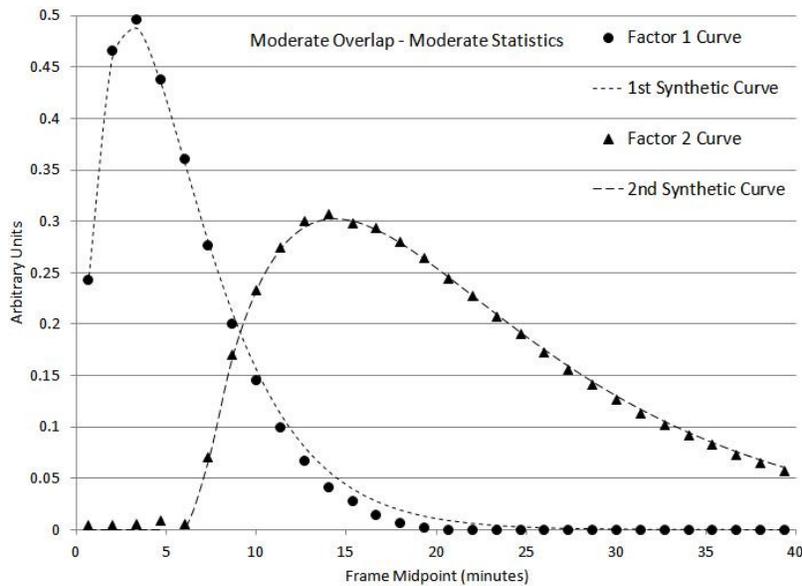
Figure 3.12: Results for the "moderate overlap" and "moderate statistics" data sets from the semi-quantitative study shown in Figure 3.7. All curves have been normalized to unity using the Euclidean norm.

3.3.3 Factor Analysis Applied to a Mouse Study

Factor images from the mouse study are shown Figure 3.13. The first two factor images correspond to the two factor curves derived in the FA study. Each factor image can be interpreted as a weighting map. Voxel values within each factor image correspond to how well the TAC derived from that voxel resembles the corresponding factor curve. In fact, the weighting values themselves are the covariances between the measured TAC and the factor curve. The third image, shown on the right in Figure 3.13, is a "residual" image. This residual image does not correspond to a factor curve itself. The image represents regions which do not resemble any combination of the two factor curves. FA effectively discards information from this image. The factor curves corresponding to these images are shown in Figure 3.14.
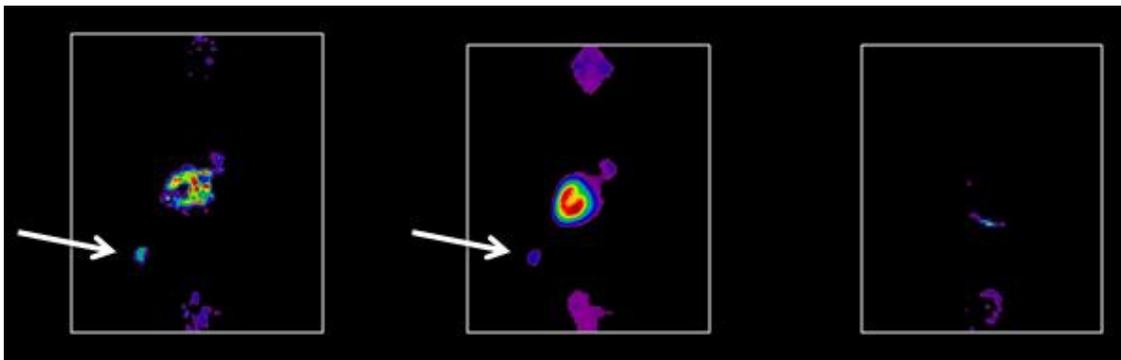
Figure 3.13: Coronal views of factor images produced from Factor Analysis of a dynamic image of tumor-bearing mouse. Vascular component (left), tissue component (middle), and residual (right). The arrow indicates the location of tumor.
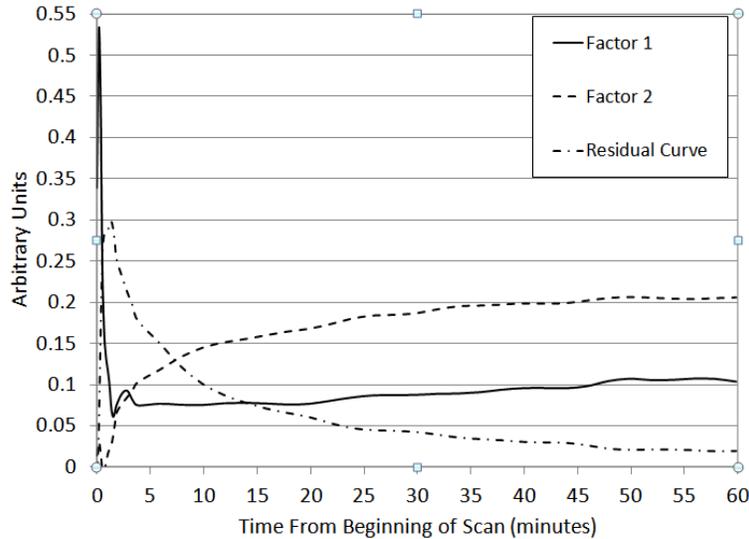
Figure 3.14: Factor curves used to derive the factor images shown in Figure 3.13. The factor images shown Figure 3.13 are weighting maps; the voxel-based weights represent the covariances between the factor curves and the voxel-derived TACs. Note that factor 1 resembles vascular kinetics while factor 2 resembles tissue kinetics. The residual curve appears to have some structure temporally, but doesn't does not contribute significantly to most voxels as seen in the residual image of Figure 3.13. All curves have been normalized to unity using the Euclidean norm.

3.4 Discussion

3.4.1 Dual-Isotope Experiment

The dual-isotope experiment provided a test for which the ideal results of FA are known and for which the results can be easily quantified. Note that, mathematically, the separation of data into two factors is the simplest scenario for FA. Furthermore, since the images consisted of only two highly localized "point sources", the statistical noise from each source is considered minimal. Therefore, the only sources of variation in the signal are from the inherent isotope decays and the instrumentation noise. The variation due to isotope decay can be model based on

simple exponential decay laws. As seen in Table 4, FA managed to recover only 45.4% of the data even when using the full data set. However, this percentage includes the "dominant" portion of the data. In other words, if the variance of a signal due to the change in kinetics is greater than the variance due to noise, FA will produce results that are more representative of the kinetics, and less influenced by the noise.

It is desirable to retain as much data as possible when using FA. However, in general, using a few factors to represent data will cause FA to effectively discard some of the data. Note, however, that the variance of this discarded data may represent the undesirable noise characteristics. Therefore, application of FA, even with a moderately low percentage of retained data, may still be feasible. In short, the percentage of data retained is not necessarily an indication of how well FA will separate data due to the fact that not all of the variance within the data stems from a true change in the signal. Conditions detrimental to FA will occur when noise is present within regions without a true dynamic signal.

Unfortunately, when applying FA to the dual-isotope data in order to derive factor curves (as opposed to principal components), meaningful results were not achieved. Additional attempts to compute factor curves were made on smaller data sets (10 of the highest-variance TACs from each of the two ROIs) using 100% of the data, but were still unsuccessful. As an alternative to using the results of PCA as an initial value in the iterative FA, measured TACs may be used. Such TACs may be determined by finding pairs of TACs such that the covariance between pairs is as close to zero as possible. This method will not only insure that meaningful factor curves are computed, but may also promote convergence since the initial values are closer to the measured curves.

3.4.2 Factor Analysis Applied to a Synthetic Data Set

Much like the dual-isotope experiment, the synthetic data set allows for a test of FA. However, in the case of the synthetic data, a signal is measured across the FOV, and not localized to a "point". Furthermore, the experiment allows for overlapping structures. As seen in Figure 3.9 through Figure 3.12, there is a very strong agreement between the synthetic curves and factor curves regardless of the noise characteristics or degree of overlap. It should be noted that even the "moderate statistics" data set consisted of a fairly robust set of events; images of this quality may not be captured in actual mice studies.

The insensitivity of FA to the spatial relations of overlapping structures is to be expected. Factor analysis is a nonparametric technique; the order in which voxel-derived TACs are arranged in FA does not change the overall statistical behavior. This nonparametric nature of FA is what differentiates it from conventional segmentation methods. The difference in kinetics between structures, as measured by covariance, is what FA relies on to separate TACs into factors.

The results of the synthetic data set are not definitive proof of the applicability of FA, but they do illustrate the potential of FA to separate structures under proper conditions regardless of the spatial arrangements. The results suggest that the factor curves provide a slight overestimate of the synthetic curves during periods of increasing activity, and vice-versa. In addition, factor 2 exhibits dynamic properties even when the corresponding synthetic curve has a value of zero during the approximate period between 0 and 5 minutes in Figure 3.10. Note that

47

the dynamics of the second factor during this period resemble that of the first synthetic curve during the same period. This suggests the interference of factor 2 by the first synthetic curve. This effect is exaggerated for the case in which there is a high degree of overlap between structures. The spatial relations between regions do not have a significant effect on the results of FA. However, the dominance of kinetics between overlapping regions may interfere with the results of FA.

The semi-quantitative study is, in effect, a study in the linearity properties of the imaging system and the FA method. The study consisted of scaling and adding independent data sets in order to synthetically create a new data set. If nonlinearities during the image reconstruction are introduced, Factor Analysis would not have been successful in recovering the original independent data.

3.4.3 Factor Analysis Applied to a Mouse Study

The dual-isotope experiment and synthetic data set provided controlled systems in which the behavior of FA results could be investigated. However, the complexities of a living subject cannot be easily simulated. A mouse study was conducted to evaluate the usefulness of FA in practice. As seen in Figure 3.13, the tumor was not entirely separated from the other structures within the mouse. Two factors were implemented in FA, resulting in the factor curves shown in Figure 3.14. It should be noted that the factor images and factor curves have no prior physiological knowledge.

The designation of the first factor curve as "vascular" and the second curve as "tissue" is purely based on interpretation. In addition, these two factor curves account for approximately 79% of the variance in the data. While this value may seem quite high compared to those found in the dual-isotope experiment (Table 4), it should be noted that the processing of the raw voxel-derived TACs was used prior to FA. This processing included smoothing and thresholding. Inclusion of a third factor (not a residual) would increase the data retention. However, inclusion of a third factor would require recalculation of the first two factors in addition to the third factor. It is unlikely that the incremental increase in data retention would justify the additional computational expense required to compute three factor images.

The results of the mouse study also support the notion that separation may be possible if the variance of data retained by FA is dominated by tracer kinetics even in the presence of moderate noise. In other words, the noise itself, nor the SNR, is indicative of the performance of FA. The performance of FA is determined by the variations of the signal compared to that of the noise. In other words, FA attempts to recreate as much of the kinetics as possible, but is unable to distinguish between kinetics due to noise versus that of the desired signal. As a consequence, the kinetics of the desired signal must be assumed to be generally prevailing before using FA.

In addition to the study shown in Figure 3.13, FA was applied to several other mice studies of various using various tracers, cell lines, tumor locations, scanning protocols, etc. Unfortunately, FA was unable to provide meaningful results from these studies. It is not quite clear why FA failed for these studies. It is possible that limited uptake within the tumors prevented FA from obtaining adequate kinetic information.

3.5   Summary

Three separate studies were performed to assess the use of FA on images from the Genisys preclinical PET system. The results from the dual-isotope experiment indicate that the noise characteristics of the system hinder the ability of FA to retain much of the data. However, the ability to retain data is not necessarily an indication of how well FA will identify and separate the prominent kinetics. The semi-quantitative study indicates the potential of FA when separating structures, even with overlapping images. The mouse study results indicate that certain kinetic properties of tumor, tissues, etc. may interfere with the ability of FA to perform a complete separation of structures. However, the results also provide insight into the kinetic properties of these structures. The work shown here indicates that FA might be applicable on the Genisys, but further quantitative studies must be performed in order to assess the optimal conditions.

4   Principal Component Analysis for Performance Metrics and Factor Analysis Limits

4.1   Introduction

4.1.1   Principal Component Analysis Primer

Principal Component Analysis (PCA) is often used as a technique for image processing due to its straightforward interpretation and convenient mathematical properties. However, for the purposes of this study, we investigate the use of PCA as a tool to characterize the system performance of a microPET system. Dynamic images of a uniform cylinder are acquired and subsequently processed using PCA. The results are interpreted in terms of uniformity; the signal-to-noise ratio (SNR); the noise distribution both spatially and temporally; and as an upper bound to Factor Analysis.

Principal Component Analysis is a method used to reduce the number of independent variables required to explain the statistical features of a given data set [13, 15, 16]. It attempts to explain as much of the statistical features as possible by using as few variables as possible. From a mathematical viewpoint, PCA exploits orthogonality between independent variables, and creates new variables. These new variables are also mutually orthogonal, and are derived from a linear combination of the original variables. Computing these linear combinations is the essence of PCA, and can be constructed as an eigenvector problem [17]. In geometrical terms, PCA performs a rigid rotation in the "data space" to maximize alignment with the most prominent features. In short, PCA attempts to compress the data so that they can be represented with as few variables as possible. This is similar to other methods such as Fourier Transforms. However,

PCA does not restrict itself to sinusoidal components; it constructs components that will maximize compression [16]. The only restriction is that these components are orthogonal. Non-orthogonal components, which allow for interdependent variables, may be computed using other methods such as Factor Analysis (see Chapter 3).

Principal Component Analysis (PCA) is a non-parametric technique; the order in which data is collected or computed is inconsequential. In addition, PCA assumes that the original variables of the data are independent. Furthermore, PCA produces a new set of variables in such a way that each new variable retains as much information as possible without redundancy. As a result, it is possible to retain most of the data features using fewer new variables than original variables. Finally, PCA does not necessarily result in components which are physically measured; the results of PCA must be properly interpreted for its physical relevance to the actual data.

Note that PCA can be performed on any symmetric matrix. These include the correlation matrix and covariance matrix. Correlation and covariance are both metrics for similarity between two signals. For example, if a data set consists of a set of time-activity curves (TACs), the correlation and covariance matrices express how similar a given TAC is to another given TAC. Correlation between two TACs, $T_i(t)$ and $T_j(t)$, as defined here, is computed by multiplying element-by-element, and then summing each product [18].

$$corr(T_i, T_j) = \sum_t T_i(t) T_j(t)$$
<div align="right">Equation 4.1</div>

Note that the correlation between two TACs may be thought of as the "inner product" or "projection" of one of the signal onto the other. Computing the covariance between two TACs is similarly performed; however, each TAC is mean-shifted prior to the element-by-element multiplication and summing. In addition, this sum is scaled by the total number of frames, n [18].

$$cov(T_i, T_j) = \frac{\sum_t (T_i(t) - \overline{T_i})(T_j(t) - \overline{T_j})}{n}$$ Equation 4.2

The mean of each TAC is computed conventionally.

$$\overline{T_i} = \frac{\sum_t T_i(t)}{n}$$ Equation 4.3

Note that the correlation between two TACs can be interpreted as similarity of the "raw" curves. Covariance is an indication of how much one curve deviates from its mean while the other deviates from its own mean. The correlation of a given TAC with itself is simply the Euclidean length of the TAC (not the area under the curve of a TAC). The covariance of a given TAC with itself is simply the variance of the TAC. Covariance of two curves may also be thought of as correlation of the same two curves after mean-shifting. These calculations can be applied to any pair of curves from the same data set.

For practical purposes, measured TACs are comprised of a true signal as well as a "residual" noise component. Under most conditions, reliable separation of the true signal from the noise is very difficult. However, if the TACs are derived from an activity source with no inherent kinetics, the true signal within each TAC is a constant, time-independent value. The variance observed within the TAC is then due entirely to noise. In other words, if each TAC is composed of a time-independent signal, S_i, and a time-varying noise component, $N_i(t)$, the measured TAC is given by

$$T_i(t) = S_i + N_i(t) \hspace{4cm} \text{Equation 4.4}$$

If the noise component, $N_i(t)$, has a mean of zero, then the true signal can be computed as the average value of the TAC.

$$S_i = \frac{\sum_t T_i(t)}{n} \hspace{4cm} \text{Equation 4.5}$$

Conversely, if one assumes that the true signal is given by the average value of the TAC, then the assumption of noise with zero-mean directly follows.

Such a data set can be measured by imaging a phantom with uniformly distributed activity. Dynamic reconstruction can be performed after histogramming with decay correction. The uniform nature of the phantom results in a true signal that is spatially uniform, while the

decay correction insures that the true signal is also temporally uniform. As a result, all voxel-derived TACs will have very similar true signals. In other words, the true signal present in each voxel-derived TAC will not depend significantly on the voxel location within the field-of-view (FOV). Ideally, all TACs derived from such a data set would be time-independent and identical. In such a case, the correlation between any two TACs would be equal to the same constant.

All of the elements of the corresponding correlation matrix, therefore, are also equal to this constant. When performing PCA on such a matrix, only one component is found. Since it is the only component, it explains 100% of the data. In practice, the noise present in each TAC will reduce this. Therefore, the percentage of data explained is a metric for the level of true signal present relative to the total signal. Due to the orthogonal nature of PCA, signals must be combined in a Euclidean (i.e. quadrature) fashion. The percentage of data explained, or retained, with one component, is therefore

$$\% \ of \ signal \ explained = {S_{total}^{2}} \Big/ {T_{total}^{2}} \qquad \text{Equation 4.6}$$

The true signal within each TAC as measured by PCA, S_i, is simply the component of the TAC that is "selected" by the first component. This selection is done simply by computing the correlation of the total TAC with the principal component, $p(t)$.

$$S_i = corr(T_i, p)$$ <div align="right">Equation 4.7</div>

The principal component computed using PCA, *p(t)*, is normalized to one, using the Euclidean norm, in order to prevent any bias. Due to the large number of TACs, computation of *p(t)* cannot be done by traditional eigenvector methods. The details of computing *p(t)* are described within Chapter 4.2.1. The selected true signal, S_i, can be thought of as a weighting coefficient. This coefficient may vary from TAC to TAC (i.e. voxel to voxel). As a result a weighting map may be computed; this map indicates the distribution of true signal within the FOV.

The total true signal present measured by PCA is found by adding these components in quadrature.

$$S_{total}^2 = \sum_t \sum_i S_i^2$$ <div align="right">Equation 4.8</div>

Note that, even though S_i, is not time dependent, it must be summed over voxel and time. The total signal measured is found by adding all measured TACs in quadrature.

$$T_{total}^2 = \sum_t \sum_i T_i(t)^2$$ <div align="right">Equation 4.9</div>

4.1.2 Correlation-Based Principal Component Analysis, Uniformity, and Signal-to-Noise
Ratio

A simplified, geometrical illustration of the principal component, when using the correlation-based method, is shown in Figure 4.1.
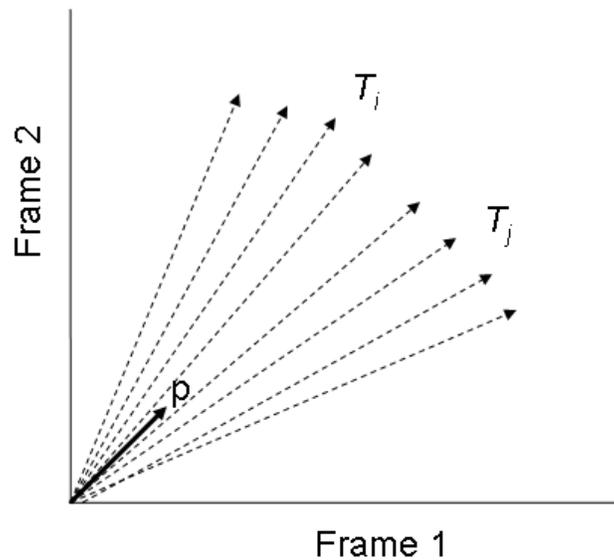
Figure 4.1: A graphical representation of correlation-based Principal Component Analysis. The principal component, *p,* is chosen such that the projections of all TACs onto the principal component are maximized. The component of a given TAC, T_i, which lies parallel to *p* is interpreted as the true signal as shown in Figure 4.9. The component orthogonal to *p* is interpreted as the noise present within the TAC. If TACs are derived from a dynamic image consisting of a uniform phantom, the first principal component may be calculated as the normalized mean of all TACs.

For this simple illustration, the entire dynamic data set consists of eight voxels reconstructed over two frames. Each TAC is represented by a vector. Since all measured TACs include positive values, all vectors lie within the upper-right quadrant. The correlation between two

57

arbitrary vectors, T_i and T_j, is given by Equation 4.1. This correlation is geometrically equivalent to the "projection" of one of these vectors onto the other. The principal component, p, is a normalized unit vector in the general direction of all other vectors. The quantity S_i (Equation 4.7) corresponds to the component of vector T_i that lies parallel to p. Stated another way, the "projection" of a vector T_i onto p is the true signal within TAC i. Note that p is 45 degrees from either axis; this is due to the fact each TAC is, ideally, a constant value. In the ideal noise-free case, all vectors would lie entirely parallel to p. However, in practice, the measured TACs include a random component within each frame. As a result, the TACs "fan" away from p. Since this orthogonal component is defined to be zero-mean, this spreading is somewhat symmetric about p. The component of each TAC which lies orthogonal to p is interpreted as representing this noisy component. The percentage shown in Equation 4.6 is essentially a measure of uniformity of all nonzero TACs. Note that, if each TAC were simply a scaled version of every other TAC, the percentage calculated would equal 100%. In other words, the image would be completely uniform over time.

Note that this percentage (Equation 4.6) involves summation of all TACs both spatially and temporally. This percentage does not quantify the distribution of signal or noise; it simply provides a metric for how much true signal is present within the total signal. All voxels are, by nature of PCA, treated as independent samples of this signal. In addition, once this percentage is calculated, rearrangement of Equation 4.4 and Equation 4.6 yields the SNR, calculated as

$$SNR = \sqrt{(\% \text{ of signal explained})\Big/(100\% - \% \text{ of signal explained})} \qquad \text{Equation 4.10}$$

$$= \sqrt{\sum_t \sum_i S_i^2 \Big/ \sum_t \sum_i N_i(t)^2}$$

Using this definition of SNR, the spatial and temporal properties of the signal are both taken into account. As indicated in Equation 4.10, S_i must be summed over voxel indices, i, as well as time points, t, even though it is a constant (i.e. independent of time).

4.1.3 Covariance-Based Principal Component Analysis and Noise Distribution

Applying PCA to the correlation of TACs provides information about the true signal within the total signal. Applying PCA to the covariance of TACs, however, provides information on the noise component itself. Note that the covariance between a pair of TACs (Equation 4.2) requires subtraction of the means of each TAC. For the simple model presented in Equation 4.4, the mean of a TAC is the true signal itself. As a result, computing the covariance between such TACs is mathematically equivalent to computing the correlation of the "residual" noise components themselves.

$$cov(T_i, T_j) = corr(N_i, N_j) \qquad \text{Equation 4.11}$$

The correlation of zero-mean noise with any other signal is, ideally, zero if the noise is in fact randomly distributed. As a result, the covariance between two different TACs is, ideally, zero. The covariance of a given TAC with itself, however, is simply the variance of the noise component under this simple model (Equation 4.4).

The corresponding covariance matrix computed using this model is, therefore, diagonal. Note, however, that not all diagonal elements are necessarily equal. These diagonal values all become equal only if one further assumes that the noise is distributed uniformly spatially and temporally. Note that, in such an ideal case, there is no single component which dominates the data; in other words, the noise component of the data is only partially explained by any given component. In total, n components will be needed to retain 100% of this data. Since there are n components, each component, including the first component, contains approximately an equal percentage of the data.

$$\% \textit{ of variance explained } = \ ^{100\%}/_n \hspace{3cm} \text{Equation 4.12}$$

In practice, the computed percentage for some components may be greater than this. This may be due to an uneven distribution of noise. Due to the nature of PCA, n orthogonal components are needed to retain all of the data. The first component, by definition, retains the most data. Therefore, the first component must retain the percentage shown in Equation 4.12 at the very least; a percentage lower than this is not mathematically possible if PCA is performed correctly. Due to the ambiguity of the principal component of this matrix, the component is not physically meaningful in itself. The percentage of data retained by this component, however, is

indicative of how uniformly the noise is distributed spatially and temporally. Furthermore, this percentage may change when varying activity levels, particularly at activity levels outside of the nominal range of the system due to nonlinearities within the imaging process.

A geometrical illustration of the principal component, *p*, using the covariance of data is shown in Figure 4.2.
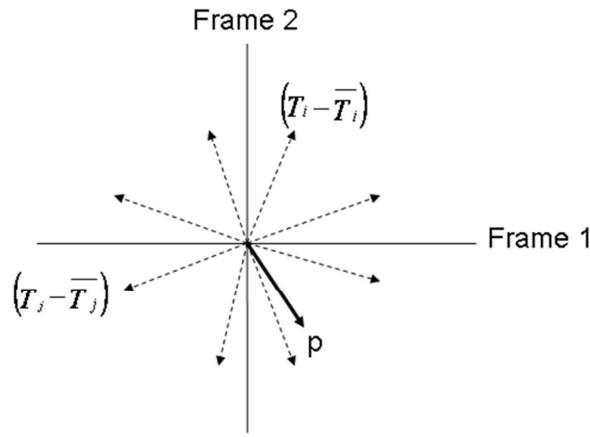
Figure 4.2: A graphical representation of covariance-based Principal Component Analysis. The vectors represent the mean-shifted TACs. For the idealized case, this amounts to only the noise components of each TAC. If noise is randomly distributed spatially and temporally, the orientation of *p* is essentially arbitrary. Although *p* itself does not have a physically meaningful interpretation, the total of the projections onto *p* can be used to calculate the percentage of variance explained, as in Equation 4.12.

Similar to Figure 4.1, the entire dynamic data set consists of just eight voxels reconstructed over two frames for simplicity. Note that all data vectors are created by shifting each TAC, T_i, by its mean, as indicated in Equation 4.2. Note that negative values may result from mean-shifting; as a result, some of the vector components may be negative unlike the components shown in Figure 4.1. Since only the noise component of the signal remains after this shifting, the vectors shown are randomly oriented in a radial fashion. Furthermore, each TAC has

essentially the same constant value in the case of a uniform cylinder; as a result, each vector shown has approximately the same norm. In the ideal case shown in Figure 4.2, there is no general direction of all vectors; the orientation of p is essentially arbitrary. Regardless of the orientation of p, the total projection of all vectors onto p is the same. This total projection, when added in quadrature, can be compared to the total norm of all vectors. This ratio results in the percentage of variance explained shown in Equation 4.12.

Note that, if noise is evenly distributed as suggested by Equation 4.12, then all subsequent components are responsible for explaining the percentage of data that is not taken into account by the first principal component. This fact will prove useful when evaluating the limits of FA, since only first components are computed for the studies described here.

It is physically possible that the noise is randomly distributed, but is not zero-mean. In such a case, the noise is biased. Dark noise resulting from the instrumentation electronics, for example, occurs even without a true signal source as may be seen on a blank scan. If so, the mean value of the noise, along with the true constant signal, will be included in the calculation of covariance (Equation 4.2). As a result, the computed covariance only retains information regarding the zero-mean component when applied to a data set with constant signal.

4.1.4 Insights from the Uniform Model

Practical imaging consists of data from uneven, time-varying activity distributions, and varying frame durations. As a result, the simple model presented here may appear irrelevant or

too naïve for useful application. However, when applying PCA to data sets of fixed frame durations or activity levels, performance characteristics of the system can be quite apparent.

For time-varying activity sources, the true signal will not be a time-independent constant. The PCA results, however, give an indication as to how much of the variation in the observed signal is due to noise. For example, suppose the percentage of explained signal (Equation 4.6), using this simple model, is 80%. When assessing the variation of a time-varying source, it can be estimated that 20% of the variance in the observed signal is due primarily to noise and not due to the actual variation of the activity concentration.

When attempting to reconstruct dynamic images, a balance must be struck between the number of frames and image quality; a large number of frames is desirable for kinetic analysis, but long duration frames are desirable for image quality. By comparing the results of PCA for various frame durations, the overall improvement in image quality, at the sacrifice of less kinetic sampling, can be quantitatively assessed.

4.1.5 Upper Bounds to Factor Analysis

Due to the close relationship between PCA and Factor Analysis (FA), the percentages computed from PCA are equivalent to those derived from FA. In fact, applying FA with only one factor will yield a solution nearly identical to that of the first component derived from PCA. As a result, the PCA results applied to this simple model give an upper bound estimate of how much data will be explained using FA. This can be generalized to situations in which more than one factor is being derived in FA.

4.1.6 Summary of Principal Component Analysis

In summary, PCA provides a quantitative method of decomposing data into basic components. These components are mutually exclusive (i.e. orthogonal), and constructed from a particular combination of the original variables from which the data was taken. The first component, by definition, retains more of the data characteristics than subsequent components. These characteristics can be measured by a number of metrics. For PCA purposes, these characteristics are measured as correlation and covariance. Both metrics lend themselves to PCA and have a straightforward interpretation. When applying PCA to the data correlation matrix, the results are interpreted as a measurement for the true signal content compared to the total signal. These results can be mathematically rearranged to determine SNR. When applying PCA to the covariance matrix, the noise characteristics are emphasized. The results from this are representative of how evenly distributed the noise is within the FOV. The simple model presented may appear too unrealistic for practical purposes. However, the results of PCA applied to a uniform cylinder can be interpreted in a number of ways. The linear nature of PCA allows the results to be applied as upper or lower bounds to nonuniform activity distributions. The characterization of noise by PCA is applicable to time-varying activity distributions. The PCA results also serve as a guideline for compromising the number of frames for image quality, or vice-versa. Finally, the strong link between PCA and FA implies the results of PCA to be an upper bound (i.e. most optimistic) for FA.

4.2   Methods

4.2.1 Alternative Computation of Principal Components Using Recursion

After reconstruction of all images, PCA was individually applied to each dynamic image set. All analysis was performed using the Interactive Data Language (IDL) version 7.0. Note that the PCA routines within IDL only allow for analysis on the covariance matrix or the correlation coefficient matrix and not on the correlation matrix. Therefore, a separate routine must be written in order to perform the desired computations. Furthermore, the matrices computed are impractically large due to the inclusion of all voxel-derived TACs within the FOV. Note that Inveon FOV contains 128 x 128 x 159 voxels by default. For a 10-frame data set, these square matrices contain $\sim 6.7 \times 10^{12}$ elements. This can be reduced by inclusion of only nonzero TACs (i.e. TACs which have at least one nonzero sample). Inclusion of empty voxels (i.e. voxels that do not have any nonzero samples) does not contribute towards these results. As a result, only nonzero TACs were included in the analysis. The number of nonzero TACs for a given data set may vary, but is $\sim 2.0 \times 10^{3}$ using the cylindrical phantom (described in Chapter 4.2.2). Figure 4.3 illustrates a subset of nonzero TACs taken from a typical Inveon data set. As can be seen in Figure 4.3, there is considerable variation in voxel values not only between TACs, but also within a given TAC.
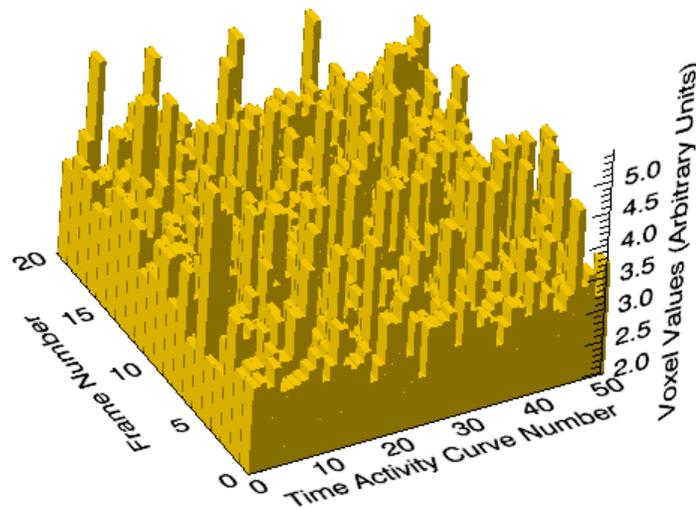
Figure 4.3: A subset of nonzero TACs taken from an Inveon 5-minute acquisition reconstructed into 20 frames using the protocol of the first study. Most data sets used in these studies consist of ~2000 nonzero TACs; only the 50 TACs with the largest norms have been shown here for visual clarity.

Even with the removal of nonzero TACs, the resulting matrices are still unwieldy. Although subsampling or smoothing can be used to reduce the number of nonzero TACs, it is believed that this processing would render the results less physically meaningful; any attempt at characterizing the system should minimize the processing of the data.

Conventional PCA methods attempt to find the eigenvalues and eigenvectors of these matrices. The eigenvector corresponding to the largest eigenvalue is, by definition, the first principal component; the relative eigenvalue indicate the percentage of data retained. As a result, it may seem necessary to find the eigenvalues and eigenvectors of these large matrices. For these purposes, however, only the first principal component is necessary; since there is only one true signal within the data, all other components are not physically meaningful. In the case of the correlation matrix, it can be shown that the first principal component, p, is given by the mean of

the TACs. A straightforward proof of this is shown in Chapter 4 Appendix A. A similar method can be used to find the first principal component of the covariance matrix.

Assuming that there are *m* nonzero TACs, this mean is given by

$$T_{avg} = \left. \sum_{i=1}^{m} T_i(t) \middle/ m \right.$$

<div align="right">Equation 4.13</div>

This is not to be confused with the mean of a given TAC. The former (Equation 4.13) is a time-dependent curve; the latter (Equation 4.3) is simply a scalar. This is normalized to unity in order to prevent bias; if not normalized, the correlation between the principal component and a given TAC will reflect not only the similarity between the two, but also the arbitrary norm of the principal component. Normalization is done in the usual way using the Euclidean norm (i.e. in quadrature fashion):

$$p = \left. T_{avg} \middle/ \left| T_{avg} \right| \right.$$

<div align="right">Equation 4.14</div>

The relationship between the mean TAC and the principal components is somewhat intuitive. Note that the principal component is designed such that the deviation of the data, when calculated about the component, is at a minimum. This minimum occurs when the deviation is

calculated about the mean of the data. (This is similar to the "DC" component of Fourier Transforms.)

Calculating the first component in this manner is mathematically correct, but does not give any of the necessary eigenvalues. In a conventional PCA computation, all of these values are needed, even if only one component is of interest. However, the eigenvalues themselves simply represent the correlation between each TAC and the corresponding normalized eigenvector. In geometrical terms, the eigenvalue is a measure of the total "projection" of each TAC onto the normalized eigenvector. Each "projection" of a TAC onto the principal component is simply the correlation shown in Equation 4.7. The total is then given by Equation 4.8. The percentage of data explained by the first principal component is then calculated based on Equation 4.6. The execution time for a typical dynamic data set is approximately 10 seconds using a 64-bit system. Sets with more frames require an execution time that is slightly higher.

Once the first component is computed by taking the mean, the additional principal components may be computed in a recursive fashion. After computing the projection of each TAC onto the first principal component, the "residual" of each TAC may be easily computed using standard vector subtraction. Once performed, the resulting data is a subset of the original data; the component of data accounted for by the first principal component is no longer presented in this subset. The second principal component, therefore, may be computed as the mean of the TACs within this subset. Additional principal components may be computed as needed using this recursive method. Conceptually, this method of successively "peeling away" principal components is similar to that of the Gram-Schmidt method of creating an orthogonal basis from a set of vectors [17]. Note that this method of computing principal components makes no

assumptions or approximations; all computations are based on the definitions of correlation (Equation 4.1) and covariance (Equation 4.2).

Covariance-based PCA can be performed in a similar manner. However, as seen in Equation 4.2, each TAC must be mean-shifted prior to the first recursion. After this, the computations for the first principal component and percentage of explained variance are identical to that of correlation-based PCA. No additional mean-shifting is required.

4.2.2 Principal Component Analysis Applied to Various Frame Durations

The first study was conducted to test the application of PCA in order to assess the quality of a given scan based on frame duration. It is intuitively clear that longer frame durations result in images with a higher SNR. Note that the increase in SNR is not necessarily proportional to the number of counts or to the square root of counts, as Poisson statistics suggest. A number of nonlinearities in the imaging process may preclude the benefit of additional counts. These may be due to detector inefficiencies, instrumentation noise, signal saturation, and reconstruction artifacts. As a result, a measurement-based method of performance characterization is used.

In order to achieve these measurements PCA was applied to a data set consisting of various scan durations and frame numbers. Data was collected from a uniform cylindrical phantom (21 mL, Ø = 25.5 mm) filled with ^{18}F-Fluorodeoxyglucose (^{18}F-FDG) a nominal activity of approximately 80 μCi. All scans were acquired on the Siemens Inveon preclinical PET scanner. Acquisitions were performed for a variety of durations: 1 minute, 2 minutes, 5 minutes, 10 minutes, and 20 minutes. Histogramming was performed using the Inveon's default

3D histogramming protocol. Reconstruction was performed using the default 3D Ordered-Subset Expectation Maximization algorithm in conjunction with Maximum A Priori (OSEM3D/MAP). The default settings were used: 2 OSEM3D iterations and 18 MAP iterations. The resulting 3D image volumes had dimensions of 128 x 128 x 159 voxels (0.80 mm x 0.80 mm x 0.86 mm). Each scan was dynamically reconstructed into 10-frame, 20-frame, and 30-frame data sets. Decay correction, deadtime correction, and detector normalization were applied for all scans. Frames were scaled such that voxel values were proportional to activity (as opposed to the total counts acquired within the frame).

The activity of 80 µCi was chosen simply because it lies within the nominal range of the scanner. Due to ^{18}F decay, the total activity within the cylinder at the beginning of each scan was slightly different. However, given the relatively long half-life ($T_{1/2}$ = 109.8 minutes) and the relatively brief scan durations, activity levels do not vary appreciably between scans. As a result, compensation for decay between scans will have little bearing on the PCA results. Attenuation correction was unavailable, and, therefore, not applied. However, it is believed that the moderate size of the cylinder precludes any appreciable attenuation artifacts.

It is possible to acquire data using only one 20-minute data set and then use subsets of the data to reconstruct the 1-minute, 2-minute, 5-minute, and 10-minute sets. However, full physical measurements were performed in order to insure that all system characteristics were taken into account.

In addition to measurements taken on the Siemens Inveon, a similar approach was taken for characterization of Sofie Bioscience's dual-headed Genisys preclinical PET system (**Figure 2.1**). Due to its own limitations, however, the scanning protocols were different from those of the

Inveon. Acquisitions on the Genisys were performed for 1-minute, 2-minute, 5-minute, 10-minute, and 20-minute data sets using a total activity of approximately 40 µCi. Histogramming was performed using 20-frame data sets. The default reconstruction was performed using an Expectation-Maximization (EM) algorithm with 100 iterations and a post-reconstruction normal smoothing filter ($\sigma = 0.35$ mm).

Once all images were reconstructed, PCA was applied to each image. Computation of the first principal component, and the corresponding percentages, was performed as described in Chapter 4.2.1. The uniformity and SNR for each dynamic set were then computed based on the methods described in Chapter 4.1.2. The noise distributions were computed based on methods described in Chapter 4.1.3

4.2.3 Principal Component Analysis Applied to Various Activity Levels

The second study was conducted on the Inveon to test the use of PCA in determining image quality across a range of activity levels. A series of dynamic images were constructed using a range of activity levels. Each dynamic set was acquired over a 20-minute period at activity levels of approximately 10 µCi, 14 µCi, 22 µCi, 31 µCi, 46 µCi, and 67 µCi of ^{18}F-FDG. All scans were performed using the same cylinder (21 mL, Ø = 25.5 mm) as that used in the varying frame duration study (Chapter 4.2.2), and positioned identically from scan to scan. Variation in activity levels was achieved using the natural decay of ^{18}F, as opposed to refilling the phantom with various activity levels. Histogramming and reconstruction protocols were

performed using the same parameters as described above. The first principal component, and corresponding percentage, was then computed as described in Chapter 4.2.1.

For the Genisys system, dynamic sets were acquired at activity levels of 6 µCi, 8.7 µCi, 12.7 µCi, 18.6 µCi, 27.3 µCi, and 40 µCi of ^{18}F-FDG. Histogramming and reconstruction protocols were performed using the Genisys' default protocols (Chapter 4.2.2). Principal Component Analysis was performed identically to that applied on the data sets acquired from the Inveon.

4.3   Results

4.3.1 Principal Component Analysis Applied to Various Frame Duration

The percentage of data retained was computed, based on the methods described in Chapter 4.2.2, for scans in which the duration was varied and for scans in which the activity levels were varied. The results from correlation-based PCA were plotted for each study separately. Figure 4.4 and Figure 4.5 respectively depict the uniformity and SNR obtained from the fixed-activity Inveon study.
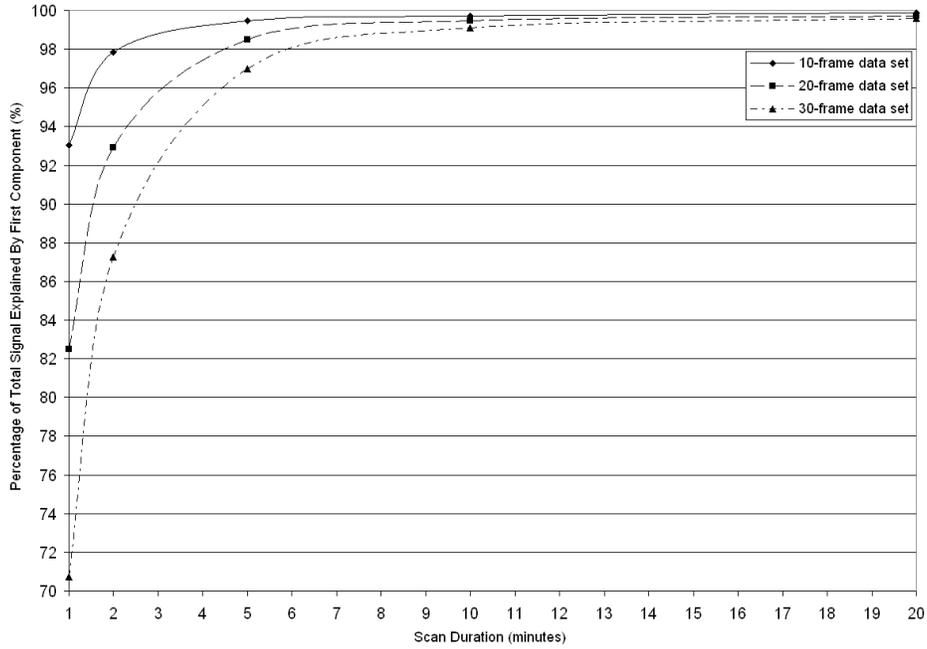
Figure 4.4: Percentage of Inveon data explained by the first component computed for each set using a total activity of approximately 80 µCi. These curves demonstrate "diminishing returns" in terms of percentage explained as scan duration is increased.
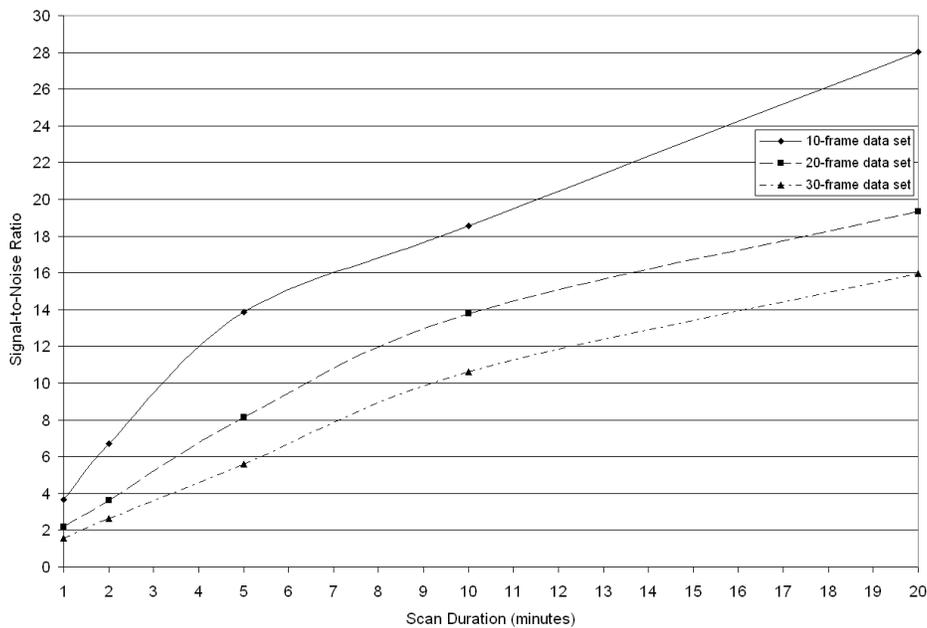
Figure 4.5: SNR calculated from PCA Inveon results shown in Figure 4.4. These curves indicate increasing benefit in SNR as scan duration is increased, regardless of the number of frames reconstructed.

73

Using covariance-based PCA, the percentage of variance explained by the first principal component was also calculated for the fixed-activity Inveon study. The percentages are shown in Table 5 through Table 7 for the 10-frame, 20-frame, and 30-frame data sets, respectively.

Scan Duration (minutes)	% of variance explained by first component
1	11.28
2	11.03
5	10.92
10	10.94
20	11.15

Table 5: The percentage of variance explained by the first component obtained from covariance-based PCA of the fixed-activity Inveon 10-frame data set. The ideal value is 10.00% based on Equation 4.12.

Scan Duration (minutes)	% of variance explained by first component
1	5.41
2	5.32
5	5.42
10	5.35
20	5.46

Table 6: The percentage of variance explained by the first component obtained from covariance-based PCA of the fixed-activity Inveon 20-frame data set. The ideal value is 5.00% based on Equation 4.12.

Scan Duration (minutes)	% of variance explained by first component
1	3.53
2	3.42
5	3.48
10	3.53
20	3.73

Table 7: The percentage of variance explained by the first component obtained from covariance-based PCA of the fixed-activity Inveon 30-frame data set. The ideal value is 3.33% based on Equation 4.12.

The results from correlation-based PCA were applied to each Genisys study as well. Figure 4.6 indicates the results obtained from the first study in which activity was kept fixed at approximately 40 μCi and scan duration varied. The SNR computed directly from these results is shown in Figure 4.7.
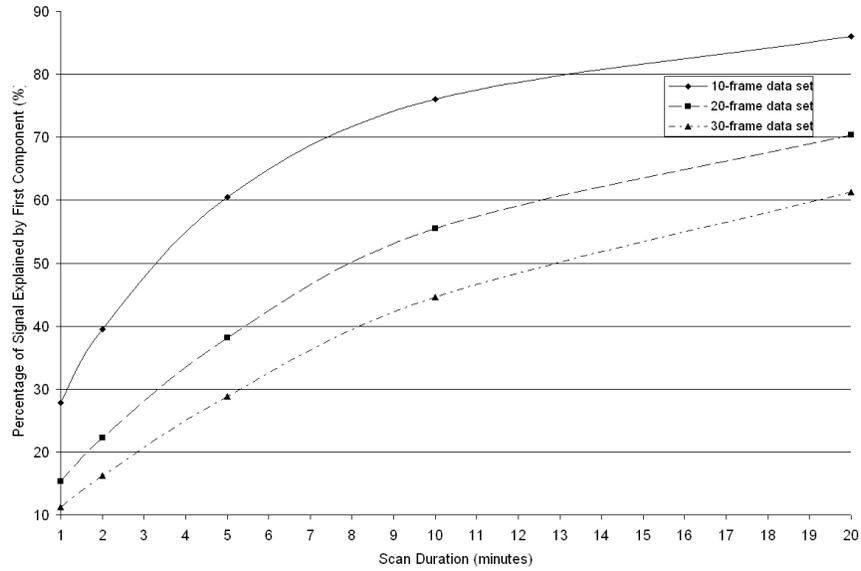
Figure 4.6: Percentage of Genisys data explained by the first component computed for each set using a total scan duration of 20 minutes for each set. These curves demonstrate "diminishing returns" in terms of percentage explained as scan duration is increased.
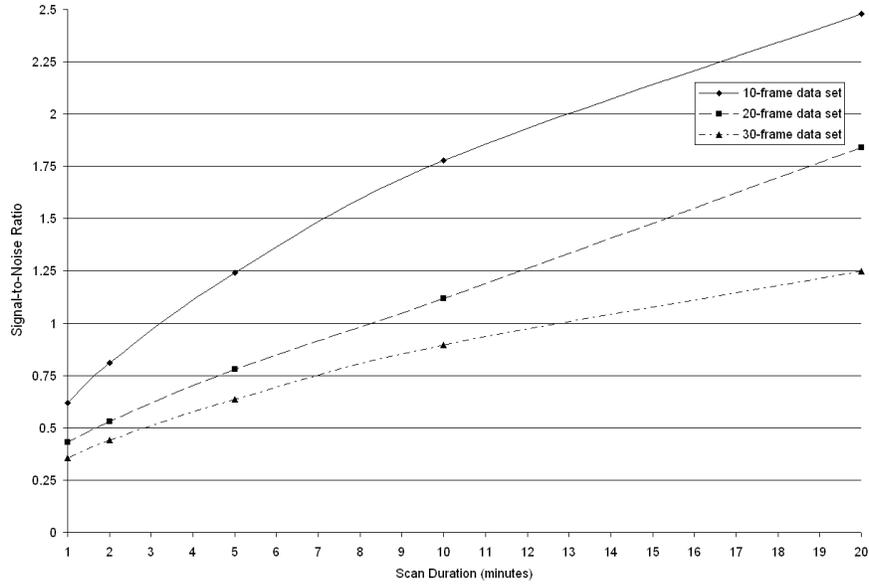
Figure 4.7: SNR calculated from PCA Genisys results shown in Figure 4.6. These curves indicate increasing benefit in SNR as activity is increased, regardless of the number of frames reconstructed. The total activity within the cylinder was approximately 40 µCi for all data sets.

Using covariance-based PCA, the percentage of variance explained for each of these data sets was computed. These are shown for the 10-frame, 20-frame, and 30-frame data sets in Table 8 through Table 10, respectively.

Scan Duration (minutes)	% of variance explained by first component
1	11.01
2	11.21
5	11.33
10	10.94
20	11.12

Table 8: The percentage of variance explained by the first component obtained from covariance-based PCA of the Genisys 10-frame data sets. The ideal value is 10.00% based on Equation 4.12.

Scan Duration (minutes)	% of variance explained by first component
1	5.29
2	5.44
5	5.39
10	5.33
20	5.34

Table 9: The percentage of variance explained by the first component obtained from covariance-based PCA of the Genisys 20-frame data sets. The ideal value is 5.00% based on Equation 4.12.

Scan Duration (minutes)	% of variance explained by first component
1	3.51
2	3.57
5	3.55
10	3.56
20	3.57

Table 10: The percentage of variance explained by the first component obtained from covariance-based PCA of the Genisys 30-frame data sets. The ideal value is 3.33% based on Equation 4.12.

4.3.2 Principal Component Analysis Applied to Various Activity Levels

The first principal component was calculated for the Inveon data sets in which the scan duration was fixed at 20 minutes, but activity levels varied. The uniformity and SNR, as measured by correlation-based PCA, are shown in Figure 4.8 and Figure 4.9 respectively.
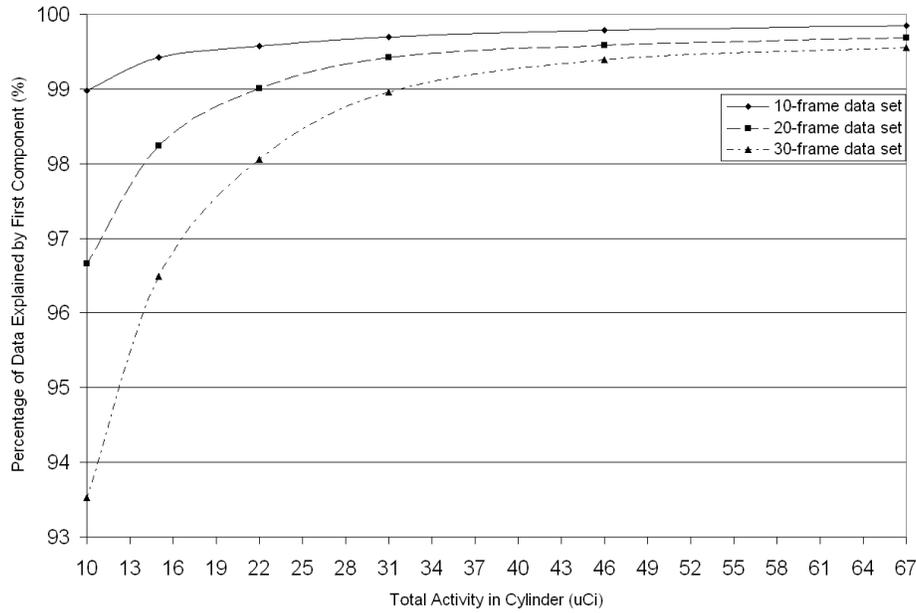
Figure 4.8: Percentage of Inveon data explained by the first component computed for each set using a total scan duration of 20 minutes for each set. These curves demonstrate "diminishing returns" in terms of percentage explained as scan duration is increased.
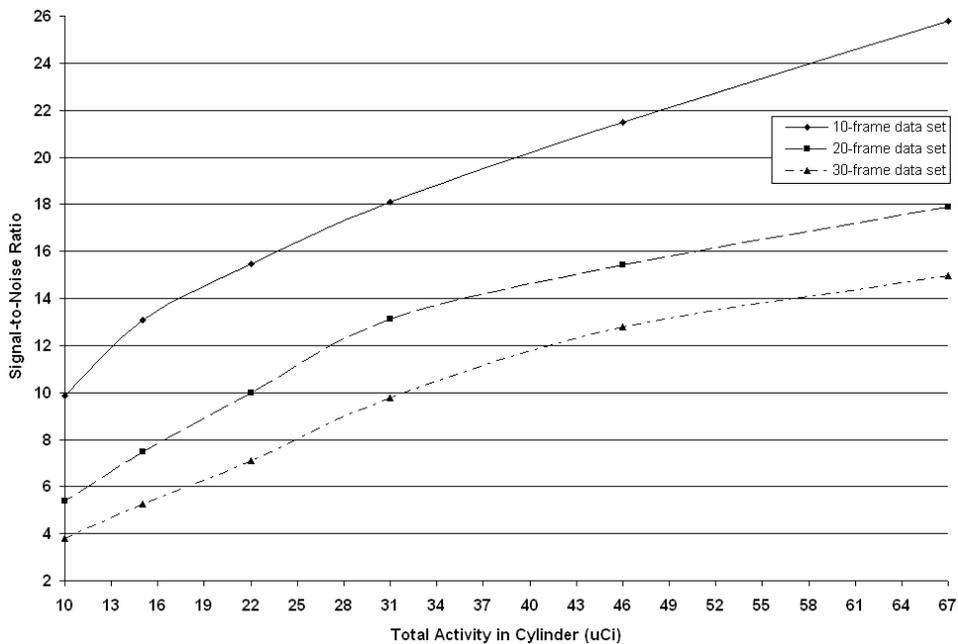
Figure 4.9: SNR calculated from PCA Inveon results shown in Figure 4.8. These curves indicate increasing benefit in SNR as activity is increased, regardless of the number of frames reconstructed.

The percentages of variance explained, using covariance-based PCA, were also computed for each data set independently. Table 11, Table 12, and Table 13 indicate these percentages for the 10-frame, 20-frame, and 30-frame data sets, respectively.

Approximate Total Activity (µCi)	% of variance explained by first component
10	11.09
15	11.36
22	11.31
31	11.43
46	11.27
67	10.09

Table 11: The percentage of variance explained by the first component obtained from covariance-based PCA of the Inveon 10-frame data sets. The ideal value is 10.00% based on Equation 4.12.

Approximate Total Activity (µCi)	% of variance explained by first component
10	5.48
15	5.74
22	5.43
31	5.17
46	5.31
67	5.27

Table 12: The percentage of variance explained by the first component obtained from covariance-based PCA of the Inveon 20-frame data sets. The ideal value is 5.00% based on Equation 4.12.

Approximate Total Activity (µCi)	% of variance explained by first component
10	3.82
15	3.72
22	3.45
31	3.45
46	3.40
67	3.52

Table 13: The percentage of variance explained by the first component obtained from covariance-based PCA of the Inveon 30-frame data sets. The ideal value is 3.33% based on Equation 4.12.

Results from the correlation-based PCA, applied to Genisys data sets with varying activity levels, are shown in Figure 4.10 and Figure 4.11.
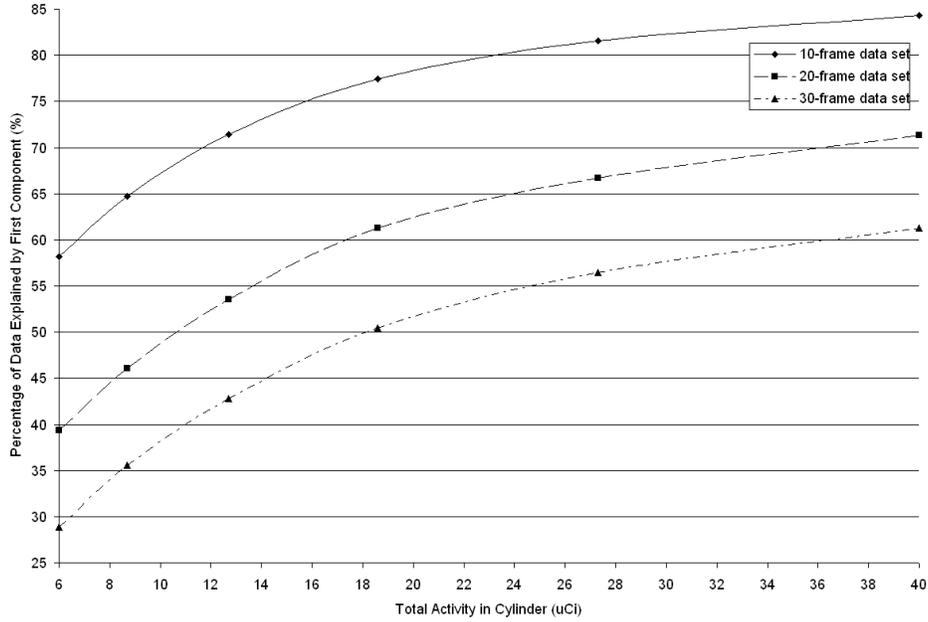
Figure 4.10: Percentage of Genisys data explained by the first component computed for each data set using a total scan duration of 20 minutes for each set. These curves demonstrate "diminishing returns" in terms of percentage explained as scan duration is increased.
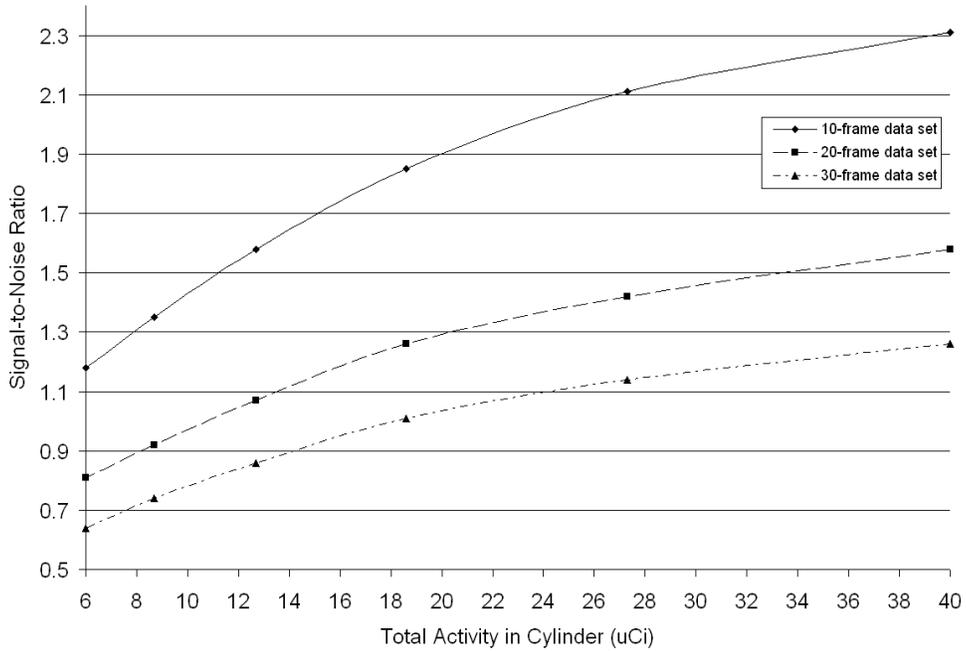
Figure 4.11: SNR calculated from PCA Genisys results shown in Figure 4.10. These curves indicate increasing benefit in SNR as activity is increased, regardless of the number of frames reconstructed.

Results using covariance-based PCA applied to the Genisys data are shown in Table 14 through

Table 16.

Approximate Total Activity (μCi)	% of variance explained by first component
6.0	11.60
8.7	11.33
12.7	11.17
18.6	11.02
27.3	11.38
40.0	11.09

Table 14: The percentage of variance explained by the first component obtained from covariance-based PCA of the Genisys 10-frame data sets. The ideal value is 10.00% based on Equation 4.12.

Approximate Total Activity (μCi)	% of variance explained by first component
6.0	5.35
8.7	5.48
12.7	5.49

18.6	5.48
27.3	5.60
40.0	5.64

Table 15: The percentage of variance explained by the first component obtained from covariance-based PCA of the Genisys 20-frame data sets. The ideal value is 5.00% based on Equation 4.12.

Approximate Total Activity (µCi)	% of variance explained by first component
6.0	3.55
8.7	3.60
12.7	3.58
18.6	3.50
27.3	3.56
40.0	3.54

Table 16: The percentage of variance explained by the first component obtained from covariance-based PCA of the Genisys 30-frame data sets. The ideal value is 3.33% based on Equation 4.12.

4.4   Discussion

4.4.1 Uniformity as Measured by the Principal Component

The results from the first Inveon study indicate that uniformity, as measured by the percentage of data explained, increases quite rapidly with scan duration as demonstrated in Figure 4.4, but quickly reach an asymptotic limit. As a result, the increase in uniformity with longer scan durations (implying longer frame durations) eventually diminishes; the uniformity is essentially independent of number of frames after a certain point. Within the 5-minute scans shown in Figure 4.4, for example, the 30-frame set and 10-frame data set have ~96% uniformity and ~99% uniformity, respectively. Only a small ~3% gain in uniformity has been achieved by tripling the frame duration. This might be further emphasized by the fact that data was acquired

using a nominal activity of approximately 80 µCi. Such behavior may not be readily seen with relatively low levels of activity.

A compromise between uniformity and frame sampling is desired; uniformity within a uniform region is, of course, preferred. A higher number of frames may also be desired for kinetic studies. The curves shown in Figure 4.4 may be used to determine a scan duration in which both needs can be quantitatively balanced. The optimal point may be defined as the scan duration at which the curve tangent (i.e. derivative) is equal to one. Prior to this point, the uniformity increases significantly with scan duration; after this point, the increase in uniformity will gradually become less appreciable. For example, the 10-frame data set in Figure 4.4 has this optimal point at a scan duration of ~2.5 minutes. This is graphically shown in Figure 4.12 where the dashed line has a slope of 1/min and is tangent to the curve at the optimal point.
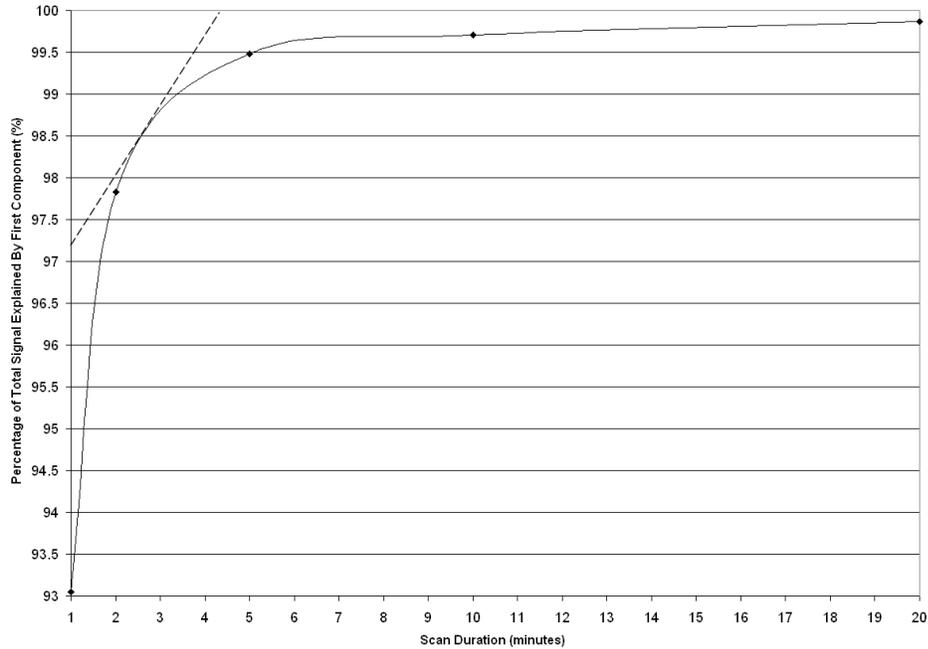
Figure 4.12: Example of finding the optimal point defined as the point of the curve at which the tangent slope (i.e. derivative) has a value of 1/min. The curve shown is the 10-frame data set rescaled from that in Figure 4.4. The 20-frame and 30-frame data sets have not been shown for visual clarity. The optimal point for this particular set occurs at a scan duration of ~2.5 minutes.

Based on Figure 4.12, a total duration of approximately 2.5 minutes should provide adequate uniformity for a 10-frame data set starting with an activity of approximately 80 µCi (using decay correction). The 20-frame and 30-frame sets have this point at scan durations of ~5 minutes and ~6 minutes, respectively. In each of the three cases, the measured uniformity is above 97%. These optimal values should be seen as a lower bound; nonuniform activity distributions which occur in practice will exhibit less uniformity, as measured by PCA. As a result the optimal points determined here will increase.

The results of the second Inveon study (Figure 4.8) exhibit similar properties. Due to the relatively long scan duration (20 minutes), the 30-frame data set has uniformity greater than 93% even at an activity level of only ~10 µCi. The uniformity dependence on activity levels is

84

surprisingly weak; activity levels beyond ~32 μCi exhibit less than 1% increase between the 30-frame and 10-frame data sets. The optimal points, as measured by activity levels, are not visible on Figure 4.8 and appear to occur at points less than 10 μCi even when using 30 frames. These points may be calculated if sets of higher frames and lower activity are reconstructed.

Very similar trends are apparent in the Genisys results. However, uniformity, as seen in Figure 4.6 and Figure 4.10, do not appear to increase as quickly as those of the Inveon, as seen in Figure 4.4 and Figure 4.8. The optimal points for uniformity appear to be greater than 20-minute scan duration for all data sets.

4.4.2 Signal-to-Noise Ratio as Measured the Principal Component

The SNR results, derived directly from these percentages (Equation 4.10) are in general agreement with what is intuitively expected. The curves shown in Figure 4.5 and Figure 4.9 indicate the increase in SNR with increasing scan duration and activity, respectively. Each curve demonstrates a somewhat abrupt rate of change. For example, the 10-frame data exhibits a decrease in the rate of change at ~5-minute scan duration. The other curves in both figures all show this, but to a lesser degree.

Optimal points which balance the needs of acceptable SNR and short scan durations can be quantified in a similar manner to that of the uniformity. In particular, the point at which the tangent slope of each curve is equal to one can be defined, again, as the optimal point. After this point, the gains in SNR are gradually diminished. For example, in Figure 4.5, the 10-frame set has an optimal point at a scan duration of ~6.5 minutes. The 20-frame and 30-frame sets both

have optimal points at scan durations of ~8.5 minutes. The 10-frame data set in Figure 4.9 indicates an optimal point at ~11 μCi. The 20-frame and 30-frame sets both appear to have an optimal point below 10 μCi.

The signal and noise are commonly assumed to follow Poisson statistics in which the standard deviation (i.e. noise) of a set of samples (counts) is given by the square root of the mean [18]. As a result, SNR is generally considered to be proportional to the square root of counts:

$$SNR \propto counts \Big/ \sqrt{counts} = \sqrt{counts} \qquad\qquad \text{Equation 4.15}$$

In this case, the counts variable refers to the total number of raw true coincidence counts detected per frame. The counts themselves can be considered to be proportional to scan duration and activity level. The general curves of each data set in Figure 4.5 and Figure 4.9 agree with the expected curve of Equation 4.15.

More quantitatively, Equation 4.15 indicates that, for a fixed scan duration, doubling the number of frames will decrease the SNR by a factor of $\sqrt{2}$. The results, shown in Figure 4.5 and Figure 4.9, are in general agreement with this. For example, Figure 4.5 indicates that the 10-frame set has an SNR of ~28 at a scan duration of 20 minutes. The corresponding 20-frame set, at the same scan duration, has an SNR of ~19. Comparison of these values yields a decrease in SNR by a factor of 1.47 which is in close agreement with the expected value of $\sqrt{2}$. However, this agreement deteriorates with lower scan duration (Figure 4.5) or lower activity levels (Figure 4.9). The Poisson model appears to underestimate the decrease in SNR when reconstructing a

given data set with fewer counts per frame. Possible reasons for the breakdown of the Poisson model include instrumentation dark noise or the inability of the iterative reconstruction algorithm to accommodate low statistic data sets.

Equation 4.15 also suggests that the SNR for a given frame is determined by its total counts. The method of how those counts were obtained, including high activity or long frame duration, is irrelevant according to the simple Poisson model. The results presented in Figure 4.5 are in general agreement with this as well. However, there is a slight, but measurable discrepancy when computing SNR using PCA. For example, Figure 4.5 indicates that a 5-minute acquisition reconstructed into 10-frames has a slightly higher SNR (~13.9) than that of a 10-minute acquisition reconstructed into 20-frames (~13.8). The discrepancy becomes more apparent when compared against the SNR of a 15-minute acquisition using 30 frames (~12.6).

Ideally, all data sets should contain approximately the same number of counts per frame. However, it is possible that the natural ^{18}F decay leads to more noise over longer acquisition periods, even with the use of decay correction. This is not supported, however, by the results of the covariance-based PCA; they strongly indicate that noise is randomly distributed both spatially and temporally. Note that the change in SNR with total counts is not as apparent in Figure 4.9 in which activity levels vary, but scan durations are fixed at 20 minutes. This suggests that, for a given number of counts per frame, there may be a modest increase in SNR when using shorter acquisitions of higher activity as opposed to longer acquisitions of lower activities.

The SNR results obtained from the Genisys data sets may be interpreted in the same manner. However, the curves shown in Figure 4.7 and Figure 4.11 do not appear to have any

point at which the derivative is equal to one; optimal points, using this criterion, are not well-defined, or possibly occur at very small scan durations or activity levels.

Poisson statistics (Equation 4.15) is generally adhered to within these results as well. For example, the 10-frame data set in Figure 4.7 indicates a SNR of ~0.62 for a 1-minute scan with 10 frames. The SNR is reduced to ~0.43 for the 20-frame data set of the same scan duration. The decrease in SNR is about 1.44, in rough agreement with the expected value of $\sqrt{2}$. The Inveon results appear to diverge from the Poisson model at low counts; however, no discernible trend was found for the behavior of the Genisys system. The rough agreement between the computed SNR and the Poisson model appears to be constant throughout the range of measured scan durations and activity levels.

As in the case of the Inveon results, the SNR is significantly different even between data sets with roughly the same number of counts per frame. For example, Figure 4.7 indicates an SNR of ~1.58 for the 10-frame data set at an activity level of approximately 12 μCi. Therefore, one would expect approximately the same SNR for the 20-frame data set at an activity level of 24 μCi. The PCA results, however, indicate an SNR of ~1.40. For the 30-frame data set at an activity of 36 μCi, the computed SNR is even smaller (~1.20).

In short, the computed SNR of the Inveon is generally an order of magnitude greater than that of the Genisys. Both systems generally follow the Poisson model when observed at fixed scan duration or fixed activity level. The model appears to be more valid for higher durations or higher activity levels within the Inveon system; this trend is not apparent within the Genisys results. None of the results, from either system, indicate an improvement in SNR greater than that of the Poisson model.

4.4.3 Noise Distribution as Measured by the Principal Component

Table 5 through Table 16 indicates the percentage of variance explained by each principal component using the covariance-based method. These results are indicative of how evenly distributed the noise is both temporally and spatially. Due to the mean-shifting of each TAC performed, the true signal itself should have no bearing on the results; only the random zero-mean component of the total noise is reflected in these results. If noise is uniformly distributed temporally and spatially, the resulting percentages are ideally given by Equation 4.12. Percentages lower than these ideal values are mathematically impossible.

Note that, in all cases, the resulting percentage is slightly higher than the ideal value. At most, the calculated percentage is 1.6% greater than the expected value (as seen in the first row of Table 14). Somewhat surprisingly, this noise distribution does not appear to be correlated with scan duration as seen in Tables 1-3 or with activity levels as seen in Tables 4-6. No discernible trend can be seen within a given set of results. However, in both studies, the 10-frame sets seen in Table 5 and Table 11 exhibit a greater deviation from the ideal result compared to the 20-frame and 30-frame data sets.

It is possible that the fewer frames themselves lead to this. Note that the covariance between any two signals is zero if either, or both, signals are zero-mean noise. However, the signals themselves must be of sufficient length due to the statistical nature of noise; only a summation of an infinitely long noisy signal is guaranteed to have a covariance of zero (Equation 4.2). The finite summations used in practice may not lead to this. Summations of shorter length,

as in the case of the 10-frame data sets, may not converge to zero. The interpretation, however, of these percentages as metrics for noise distribution, is not affected. The results, shown in Table 5-Table 16, strongly suggest the nearly uniform distribution of zero-mean noise temporally and spatially. In addition, these noise characteristics appear to be independent of scan duration and activity level. Extremely short frames or low activity levels, beyond those used for these studies, might exhibit more coherent noise properties.

If so, this implies that the PCA results are rather independent of the number of nonzero TACs measured. In other words, the actual dimensions of the uniform phantom used in the experiment are somewhat irrelevant to the PCA results. The same appears to be true for the actual placement of the phantom within the FOV. However, obtaining an adequate number of TACs to supply sufficient statistics will still be necessary. Therefore, the uniform phantom must not be too small.

4.4.4 Upper Bounds to Factor Analysis

Due to the close relationship between PCA and FA, the results presented here may be used as a guideline for determining how much data is retained when performing FA. Note that uniform activity distribution, for a given set of TACs, is the simplest case in which FA can be performed; the lack of any inherent kinetics implies that only the variation in signal due to noise must be accommodated by the derived factors. In the ideal case, all TACs are scalar multiplies of each other (i.e. no noise). If so, only one principal component would explain 100% of the data. If using FA, only one factor would explain 100% of the data. However, due to the inability of PCA

or FA to distinguish variance from noise from that of kinetics, any level of noise present will reduce this. Furthermore, the results of PCA are designed such that each subsequent component retains as much of the data as possible without overlapping with that of previously derived components. In short, it is mathematically impossible to retain more data using a method other than PCA for a given number of parameters. Therefore, the percentages calculated from PCA should be seen as an upper bound (i.e. most optimistic) for those in FA. The actual amount of data retained by FA will be determined by the measured kinetics, number of iterations, and how sensitive to noise the particular FA algorithm is.

An upper bound to FA using one factor is simply equal to that of PCA. For example, if applying FA to a uniform distribution with a total of ~80 μCi acquired over 1 minute and reconstructed using 30 frames, Figure 4.4 gives the upper bound of approximately 71%. In the case of a uniform activity distribution, this first component accommodates the true signal itself; presumably, all subsequent components retain information of the noise only. Table 5 - Table 16 strongly suggests that the noise is uniformly distributed both spatially and temporally. As a result, the percentages of data explained by these subsequent components are approximately equal. Therefore, the remaining 29 % of the data is equally explained by the remaining 29 components. Note that the total number of components, when using PCA (or number of factors when using FA) is equal to the number of frames for a given data set. Therefore, the second component explains approximately 1 % of the data. The orthogonal nature of PCA allows for calculation of the total data explained as the sum these two components. Therefore the approximate percentage of data retained by two components is approximately 72%. This additive property may not be applicable to FA due to the possibility of interdependent (i.e. nonorthogonal) variables. As a result, this sum should be considered as an upper limit only. For

this particular example, percentages for additional components may be computed by adding 1.0% for each subsequent component. Inclusion of all 30 components must result in a total of 100% of the data.

This approximation for percentage explained becomes less valid for a larger number of components. The usefulness of PCA and FA lies in the fact that the majority of the data is compressed with a number of components that is small relative to the number of total components. In addition, FA employed in practice attempts to categorize various signals into classes with common kinetics. These signals may come from regions with different numbers of TACs or even regions of uneven activity distribution. As a result, FA attempts to explain variations not limited just to variations due to noise. The inclusion of these kinetics will only decrease the percentage of data retained relative to these estimates.

4.5   Summary

Application of Principal Component Analysis as a measure of system performance of the Siemens Inveon and Sofie Biosciences Genisys preclinical PET systems were assessed using two studies. The results of correlation-based PCA and covariance-based PCA are interpreted in terms of uniformity, SNR, noise distribution, and an upper bound for possible FA. Uniformity was defined in terms of percentages of data retained by the first component obtained from correlation-based PCA. The PCA methods were applied to data sets in which activity was held fixed and frame duration varied, and vice-versa.

Uniformity within the Inveon was found to increase quickly for low counts, but soon reached a point at which increases in uniformity were less appreciable. The corresponding Genisys results also showed increased uniformity with increasing counts, but did not exhibit the rapid growth found within the Inveon system.

The SNR of each dynamic set was computed from the uniformity measurements under the assumption of zero-mean noise. The SNR was observed to obey Poisson statistics fairly well. Agreement with Poisson statistics increased for higher count levels within the Inveon results, and generally remained constant for the Genisys. The SNR of most Inveon data sets were an order of magnitude greater than that of the Genisys.

Noise distribution was quantified using the percentage of data retained obtained from covariance-based PCA. The results from these analyses strongly suggest a nearly uniform distribution of noise spatially within both systems. This allows for a simple estimation of the upper limit of percentage retained when using FA regardless of the details of the algorithm and kinetics.

Further work with this method can be used to investigate these properties in less desirable conditions such as very short frame durations or low activity levels. It is believed that larger discrepancies will result within these ranges. In addition, the method of computing principal components, using the correlation-based method, can be extended such that all subsequent components can be explicitly computed in a recursive fashion as described in Chapter 4.2.1. Percentages retained by the second component can then be quantified and compared against the ideal case of a uniform noise distribution. Finally, the methods described are not specific to PET systems and may also be applied to other time-varying imaging modalities.

4.6   Appendix A: Derivation of the Mean as the First Principal Component

We wish to prove that the mean of a given set of TACs is equivalent to finding the first principal component. For the proof presented here, the matrix representations, eigenvectors, and eigenvalues involved in PCA [15] are not used; instead each TAC is represented as a simple time-dependent function, $T_i(t)$. The straightforward calculus-based technique for minimizing functions is used. It is nearly identically to that used to find the best-fitting curves in linear regression [18]. This method is analytic and is facilitated by the fact that the principal components follow a least-squares criterion. This criterion happens to be especially convenient for minimization purposes. In fact, PCA can be thought of as a least-squares problem with multiple, independent parameters. For our purposes, the problem is further simplified by the fact that no constraints are used on these parameters.

The quantity of interest is the time-dependent principal component, p. Assuming there are n frames, and m nonzero TACs, we wish to find all of the n values within p. For notational simplicity, the value of the principal component at a given time t is denoted as p_t as opposed to $p(t)$. Each element within $p(t)$ is a continuous variable independent from all other elements. By definition, the principal component maximizes the percentage of data explained; in other words, the principal component follows the usual least-squares criterion by minimizing the total residual error, χ^2.

$$\chi^2 = \sum_i (T_i - p)^2 \qquad\qquad \text{Equation 4.16}$$

This equation subtracts two curves (i.e. vectors), at all time points (i.e. elements), and then adds the differences in quadrature. However this can be expanded in quadrature fashion to show the subtraction between individual time points within each curve. This necessarily assumes independence between all curves at all time points, as is the nature of PCA:

$$\chi^2 = \sum_i \sum_t (T_i(t) - p)^2 \qquad\qquad \text{Equation 4.17}$$

We wish to find all values of p_t such that χ^2 is minimized. This is done using the usual calculus-based method of taking derivatives. Note that in Equation 4.17, the variable t is discrete, but all the elements of p_t are continuous. Also note that the total residual error can be made arbitrary large by making any, or all, variables p_t arbitrary large; in other words, χ^2 does not have an absolute or local maximum.

Therefore, the usual method setting derivatives to zero and solving for p_t will suffice to find a minimum. Since there is only one such minimum, these values of p_t will define the absolute minimum, which is the quantity in question. To find this minimum, we can use the usual calculus-based method of setting the derivative of χ^2, with respect to a particular time point p_k, equal to zero:

$$\frac{\partial \chi^2}{\partial p_k} = 0 \qquad\qquad \text{Equation 4.18}$$

Equation 4.18 must be zero for all values of k. Note that, in general, this condition is not enough to guarantee a local minima or maxima. Additional conditions involving the second derivatives and cross derivatives must also be used. Referring to Equation 4.17, we can see that all second derivatives and cross derivatives are identically zero. Therefore, these additional conditions for finding a local maxima or local minima are automatically satisfied. This convenient property makes the least-squares criteria extremely versatile.

Referring again to Equation 4.17 and taking the derivative with respect to p_k yields:

$$\frac{\partial \chi^2}{\partial p_k} = -2 \sum_i (T_i(k) - p_k) \qquad\qquad \text{Equation 4.19}$$

Setting Equation 4.19 equal to Equation 4.18, and solving for p_k results in

$$p_k = \frac{\sum_i T_i(k)}{n} \qquad\qquad \text{Equation 4.20}$$

where n is the number of time points. Equation 4.20 indicates that, for a particular time point k, the principal component element p_k is found by averaging the values of all TACs for that same time point.

In short, the first principal component for correlation-based PCA can be found by averaging all nonzero TACs. For covariance-based PCA, a nearly identically argument can be made. However, all TACs must be mean-shifted prior to the averaging. In other words, $T_i(k)$ must be substituted with the mean-shifted version (Equation 4.2). A separate principal component must be computed for every dynamic data set. In addition, the correlation-based and covariance-based methods each produced their own principal components. After computation of the components, the components are normalized in quadrature fashion as shown in Equation 4.14.

5   Multimodality Coregistration Using Rigid Body Transformations

5.1   Introduction

5.1.1 Rigid Body Transformations for Coregistration

Multimodality coregistration for mice imaging has been employed using a variety of methods. Some methods rely on landmarks or markers, precompiled atlases, or statistical properties of the pair of images being fused. In each case, the transformation is reliant on a fundamental assumption regarding the mice or voxel value distributions. These methods may fail when applied to anatomically unusual specimens, noisy images, or even modalities for which the assumed properties do not apply. Futhermore, the actual transformation employed for coregistration between a given pair of images may differ from pair to pair. In other words, the transformation being applied may not be consistent even if applied to pairs of images taken from the same modality. Certain methods may also alter the information from a given image for the sake of fusion thereby decreasing its quantitative value. Even if visually agreeable, it is difficult to determine how accurately the coregistration was performed. A coregistration method which does not rely on any assumptions of anatomy or on the images themselves, yet retains all the quantitative value of the images, would be highly useful.

Note that coregistration can be mathematically considered as a problem involving the alignment of two different coordinate systems. Due to the static positioning of mice during image acquisition, the mice can be effectively treated as a rigid body. As a result, coregistration of mice can be treated as a rigid body transformation between two coordinate systems. This

approach has been recognized by Ji et. al. [19]. The problem of rigid body transformation has also been used in motion correction for PET brain imaging [20, 21, 22]. The motion correction problem involves multiple frames from a single imaging system; the coregistration problem involves multiple modalities of varying image sizes. However, the fundamental concepts and algorithms are applicable in both due to the rigid body nature of the problem. In other words, motion of a rigid body and coregistration of a static subject are mathematically equivalent problems. In the case of motion correction, an external motion-tracking device may be used to provide a transformation from one image to the next. However, no such device is currently in place for the coregistration problem. The challenge lies in how to derive a transformation which puts each image into a common frame of reference. This frame of reference can be arbitrarly chosen. However, for convenience, the coordinate system of one modality is chosen to be the common frame of reference itself.

A qualitative illustration of the rigid body coregistration problem is shown in Figure 5.1. As shown in Figure 5.1, each modality, or imaging system, may be thought of as having its own inherent coordinate system. For visual purposes, only a 2-dimensional (2D) depiction is shown.
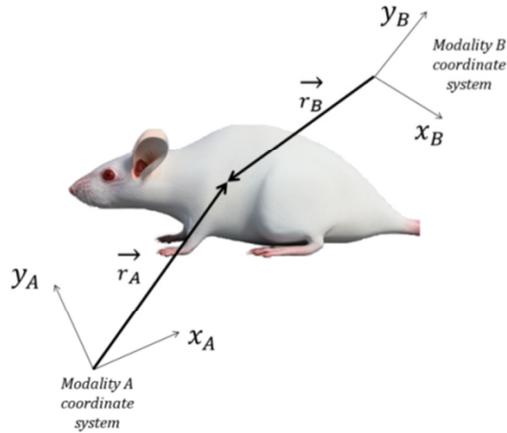
Figure 5.1: Graphical illustration of the rigid body coregistration problem. The coordinate systems of two modalities, A and B, are translated and rotated arbitrarily with respect to one another. The rigid body coregistration problem consists of finding the optimal rotation and translation such that alignment between the two systems is as close as possible. For simplicity, only a 2-dimensional view is shown. In practice, the imaging and coregistration are performed in all three dimensions.

In practice, the coregistration problem applies to all three spatial dimensions (3D) simultaneously. These coordinate systems are independent, and rotated and shifted relative to each other in an arbitrary fashion. In short, there is no universally agreed upon convention for defining each of the three axes relative to the subject. For each system, a given point within the subject is localized by a system-specific coordinate. The coordinates $\vec{r_A}$ and $\vec{r_B}$ represent the coordinates to the same point from modality A and B, respectively. The coordinates, themselves, however, are not the same due to the fact that there are measured relative to different coordinate systems. The coordinates are measured in terms of physical distance (i.e. mm); they do not indicate voxel indices. The origin of each coordinate system is chosen to lie somewhere within the field-of-view (FOV) within each system. The choice of where in the FOV is arbitrary.

The problem of interest is to compute a rigid body transformation such that the two coordinates systems align as closely as possible. By definition of rigid body transformation, the transformation may only consist of a rotation and translation; other possible transformations such as dilation, shearing, etc. are not used. However, instead of transforming the coordinate systems (i.e. "passive transformations"), the coordinates themselves can be transformed (i.e. "active transforms"). Mathematically, both concepts are essentially equivalent. For a given passive transformation, the corresponding active transformation is given by the inverse. For computational reasons, it is easier to work with the active transformations; the coordinates within a given modality, A, are rigidly transformed to match that of B. The problem of coregistration, therefore, is to find some transformation T, such that

$$\vec{r_B} = T\left(\vec{r_A}\right)$$
Equation 5.1

This transformation applies equally to all points within the subject; no adjustments to the transformation are made based on location within the subject. This principal is a defining characteristic of the rigid body transformation.

5.1.2 Computation of the Rigid Body Transformation

As mentioned, calculating T involves two pieces: calculating a 3D rotation and calculating a 3D translation. Taken separately, each piece is manageable. However, when taken

101

simultaneously, the problem is nontrivial, particularly when an analytic solution is desired. The general mathematical problem has been addressed by Horn using a least-squares criterion [23, 24]. The resulting rotational piece is expressed in terms of a quaternion, q. Although a more familiar rotation matrix can be computed, the quaternion-based method is computationally much more straightforward. Details of how a quaternion expresses rotation and its analogy to the rotation matrix can be found in Kuipers [25]. The translational piece is expressed as a standard 3D vector $\vec{t}$. In short, $\vec{r_A}$ and $\vec{r_B}$ are related by

$$\vec{r_B} = T\left(\vec{r_A}\right) = \left(q\,\vec{r_A}\,q^*\right) + \vec{t}$$

<div align="right">Equation 5.2</div>

By convention, the coordinates are first rotated using the quantity in parenthesis in Equation 5.2. (The quantity q^* is the complex conjugate of q.) After rotation, a linear shift is applied using the translation vector. The rotation and translation each provide three degrees of freedom. In other words, the transformation consists of six independent values to be determined. All other values can be computed from these six.

5.1.3 Bed Apparatus Used for Rigid Body Coregistration

The success of this coregistration method rests almost entirely on how immobile the subject is relative to the bed. Even minor repositioning of the subject between scans may affect the accuracy of the coregistration. In addition, the bed itself must be rigidly placed. In other words, the placement of the bed within the (FOV) for each system must be reproducible. For

these purposes, beds adopted for use with the Sofie Biosciences Genisys G4 preclinical PET scanner were employed. The bed dimensions and tracks were customized to fit both the Inveon and microCT. As a result, the subject can be imaged on all three systems without having to be removed from the bed. A photograph of the bed apparatus used within the Siemens Inveon preclinical PET system is shown in Figure 5.2. The apparatus is also designed to accommodate oxygenation, heating, and sedation of the subject while in the bed during acquisition (not shown).
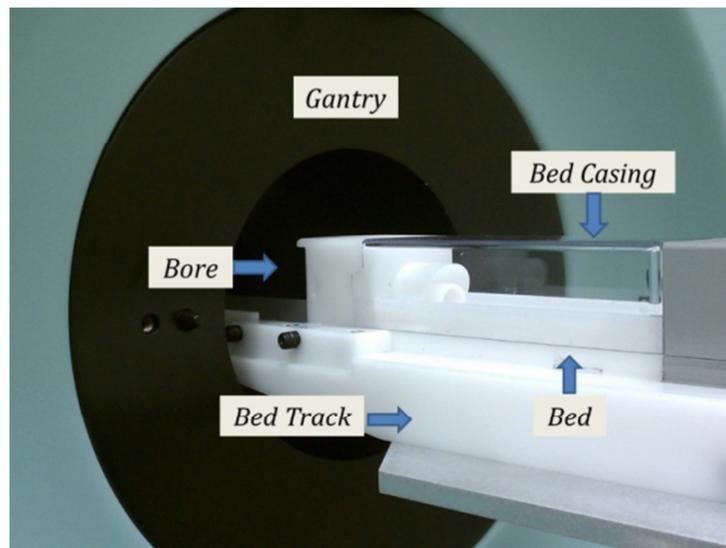
Figure 5.2: Bed apparatus used within the Siemens Inveon preclinical PET system. For visual clarity, no subject is shown, and the bed is only partially inserted into the bore. In practice, the bed is fully inserted into the bore so that the subject within the bed lies within the FOV. Identical bed tracks are fitted to each imaging system allowing the bed to be inserted into each system without having to reposition the subject between scans.

5.1.4 Coregistration Calibration Measurements

As stated, this coregistration method does not rely on landmarks or statistical properties to compute a transformation. Instead, it uses a set of calibration measurements acquired independently of any image. These measurements require the same set of points within the bed to be localized by two different systems. The concept is illustrated in Figure 5.3.
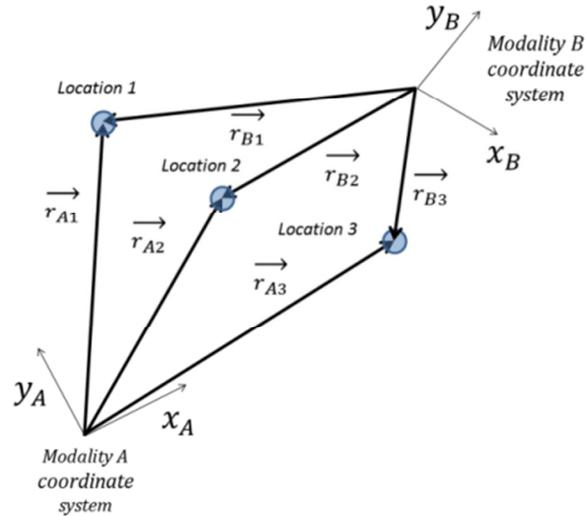
Figure 5.3: Illustration of calibration data used by Horn's method for computing the rigid body transformation. Three random points are chosen to be localized within each FOV. The points themselves are stationary relative to the bed. Mathematically, at least three noncollinear points are needed to compute a transformation.

As seen in Figure 5.3, the calibration data consists of the coordinates $\vec{r_{A1}}$, $\vec{r_{A2}}$, and $\vec{r_{A3}}$ measured in Modality A and similarly for Modality B. Note that $\vec{r_{Ai}}$ and $\vec{r_{Bi}}$ are different coordinates (regardless of the value of i), but are measuring the location of the same physical point. The choice of these points is arbitrary. However, at least three points much be used to calculate a transformation. In addition, at least three of these points must be noncollinear. (Note

that if points are noncollinear in one system, then they are necessarily noncollinear in any other system after a rigid transform.)

The localization is done in terms of physical coordinates (measured in distance), not in terms of voxel indices, voxel values, etc. This data may be thought of as calibration data for the algorithm; the derived transformation attempts to rigidly align this data as much as possible in the least-squares sense. Mathematically, there is only one such transformation for a given set of data. Due to the rigid nature of the subject, this same transformation can be applied to any pair of images from the same pair of systems as long as the calibration data and subject images were captured using the same bed positioning and axes convention. In other words, if the calibration data and subject images are acquired using different bed positioning or axes convention, the transformation is invalid. Repeated measurements of the calibration data are not necessary if no changes have been made to bed positioning or axes convention regardless of the protocols used for acquiring subject images. This consistent "cookie-cutter" approach to the coregistration is one of its key advantages. In short, coregistration is being performed not on the subject per se, but on the FOVs between the two systems.

5.1.5 Horn's Method for Computing a Rigid Body Transformation

The method of computing a rigid body transformation, simply referred to as Horn's method has been proposed and utilized in a number of applications [23, 24]. The method is well-established and has been used for coregistration between microSPECT and microCT [19], amongst other applications. The algorithm is a least-squares method which utilizes three or more

reference points. The physical coordinates of these reference points are known within each coordinate system. These reference points, not the images themselves, are used by Horn's method to compute a transformation between the coordinate systems.

The computation first shifts all coordinates within each coordinate system relative to their respective centroids. This has the effect of creating an origin for both coordinate systems which coincide with each other; in other words, the effect of translation between the two coordinate systems is removed without having to solve for it explicitly. The problem is then reduced to finding an orthogonal transformation (i.e. rigid rotation) between the two sets of redefined coordinates. This is done through a somewhat lengthy series of vector computations, at the end of which an eigenvector must be computed for a 4 x 4 matrix. Note that this is not a spatial transformation matrix (which may incorporate rotation and translation together). The 4-element eigenvector which corresponds to the largest eigenvalue of this matrix is, however, the sought after quaternion q (Equation 5.2). The translation is subsequently found using a 3D vector subtraction. The use of the maximum eigenvalue is conceptually very similar to the linear algebra approach taken in Principal Component Analysis (Chapter 4). In short, the eigenvectors may be thought of as representing axes of symmetries; their corresponding eigenvalues indicate their degree of symmetry.

The quaternion, much like a rotation matrix, specifies instructions about the rotation. The quaternion produced by the method can describe any orthogonal spatial transformation. Once this is acquired, the inverse of this rotation is applied to the original coordinates. From this, a simple subtraction yields the desired translation vector. The final transformation given by Horn's method, therefore, yields a quaternion and vector used for rotation and translation, respectively. Aside from the assumption of a rigid body transformation, no other assumptions or

106

approximations are used in the calculation. In addition, the reference points used in the calculation can constitute any set of three, or more, noncollinear points. No restrictions are imposed on the distances between points, orientation, etc. Implementation for Horn's method was performed using the Interactive Data Language (IDL) version 7.0.

Simulated studies have shown that Horn's method does, in fact, recover exact coordinates (within floating-point error) for a known, user-defined transformation. Alternatively, the transformation itself can be recovered using user-defined coordinates. In practice, of course, the transformation is not user-defined, but is prescribed by the systems being coregistered. Therefore, an investigation of this method involves quantifying how accurately the calibration itself can be measured. In short, Horn's algorithm itself is not being tested. Under the assumption of a rigid body transformation, the method provides highly accurate transformations. How well the calibration data agrees with this assumption is the point in question.

5.1.6 Least-Squares Metric for Rigid Body Transformations

As stated, Horn's method employs a least-squares criterion for determining the rigid transformation. Regardless of how rigid the data is, Horn's method [23, 24] always provides the best fitting rigid body transformation in the least-squares sense. Quantitatively, it finds the transformation, T, such that the total residual error squared is minimized. In terms of Equation 5.2 and Figure 5.3, it minimizes the quantity χ^2:

$$\chi^2 = \sum_i \left| \overrightarrow{r_{Bi}} - T\left(\overrightarrow{r_{Ai}}\right) \right|^2 \qquad \text{Equation 5.3}$$

Equation 5.3 simply finds the total residual error squared for all points before and after a given transformation. The squared quantity within bars is simply the distance (i.e. norm) between the location of a point in system B and the same point measured in system A after transforming to the coordinates in system B. Note that the total error is in units of length-squared. It may be interpreted as a goodness-of-fit parameter [18], or as the variance of a measurement of distance. For ideal, perfectly fitting data, this value will be zero regardless of which points were used in the calibration data. However, practical limitations of how well the calibration data can be measured will only increase this error.

Although the quantity shown in Equation 5.3 has physical units, it should not be interpreted in an absolute manner. In other words, if the calibration measurements are repeated using another set of randomly selected reference points, this value may change. The calibration with the larger value is not necessarily inferior. The apparent difference may be due to inaccurate localization. However, this discrepancy may simply be due to the arbitrary choice of locations used in the calibration; locations further away from the origin within a FOV may increase this quantity even if Equation 5.3 is minimized. Furthermore, including additional calibration points (beyond the requisite three) may improve the calibration, but may also increase this quantity. In short, the minimization of this quantity is of interest, not the value itself.

In the example described, coordinates from system A are converted to that of system B as seen in Equation 5.2. The inverse of this may be easily calculated by performing the inverse of the rotation and translation in the reverse order.

$$\vec{r_A} \cong T^{-1}\left(\vec{r_B}\right) = q^*\left(\vec{r_B} - \vec{t}\right)q \qquad\qquad \text{Equation 5.4}$$

Equation 5.4 first translates the coordinate from system B using the opposite shift of that computed; it then rotates the result in an opposite manner from that computed. Again, for ideal data, the location of any given point in system A can be recovered exactly using the transform and coordinate in system B. In short, there are four transformations that can be used to coregister any given pair of images. The results may differ slightly for all four cases when using less-than-ideal calibration data.

In practice, the limitations of measured data, however, indicate the recovery shown in Equation 5.4 is only a very good approximation. This limitation stems from the discretization of voxels and, therefore, coordinates. In addition, measurements taken for the calibration data may have their own errors. The errors acquired during the calibration measurements propagate into the coregistration itself. Note that the same set of points, relative to the bed, must be measured in both coordinate systems. However, the limitations of this measurement may preclude localization of exactly the same points as seen in Figure 5.4.
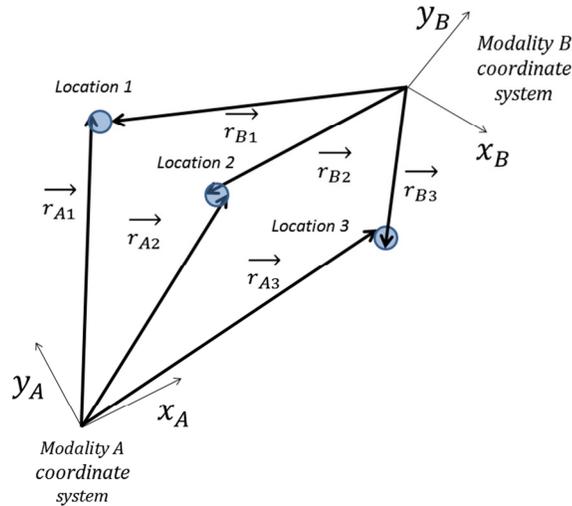
Figure 5.4: Graphical illustration of calibration data used for computing the rigid body transformation. Unlike what is shown in Figure 5.3, the same three points are not localized by both systems. The error in the measurement will degrade the accuracy of the coregistration.

For our purposes, the calibration data was acquired using a spherical phantom imaged in both systems in various locations. The various placement of the phantom is shown in blue in Figure 5.3 and Figure 5.4. The details of acquiring the calibration data are discussed in Chapter 5.2.2.

5.1.7   Summary of the Rigid Body Coregistration Problem

In summary, the problem of coregistration is inherently a spatial problem, as opposed to a statistical problem. As a result, a suitable solution is inherently spatial and makes no assumptions about voxel values, distributions, anatomy, or landmarks. Coregistration of mice imaging and motion correction of brain images share the property of a rigid body transformation. Horn's method provides a highly accurate method of calculating a giving rigid body transformation

using a straightforward analytical method. It also uses the intuitive least-squares criterion for deriving a rigid body transformation. In addition to the transformation, a metric can be computed using the least-squares criterion. Although not absolute, the metric provides a way of gauging how well the given transformation applies to the measured calibration data. Due to the reproducible nature of rigid body transformations, this computed transformation can be applied to any pair of images assuming that the bed position and axes orientations have remained consistent. Investigation into this method does not involve testing the algorithm itself; instead, the investigations involve determining how well the method can be applied under practical imaging conditions.

The experiments performed attempt to quantify how well the imaging systems correspond to rigidly transformed coordinate systems. These experiments were first measured without a subject in order to determine any errors originating from the bed apparatus or scanners themselves. Once these errors were found to be minimal, calibration data was taken to compute the transformation between them. This transformation was applied to the calibration data itself as well as to mice images to visually assess the quality of coregistration.

5.2   Methods

5.2.1 Determining Positioning Repeatability

As mentioned, the success of the coregistration algorithm depends heavily on how consistent positioning of each bed, relative to the FOV, is within each scanner. As a result, we

must quantify this consistency (or lack thereof) in order to assure that a rigid body transformation is an applicable method of coregistration. The measurement of position within the FOV is performed relative to an arbitrary origin with each system's FOV. The variance of these measurements, however, is an absolute metric for how reproducible positioning with the FOV is. The scanners used in these experiments consisted of the Siemens microCT II, Siemens Inveon preclinical PET scanner, and Sofie Biosciences Genisys G4 preclinical PET scanner. The microCT and Inveon were both fitted with the bed apparatus shown in Figure 5.2; the apparatus was not required for the Genisys G4 due to the fact that the inherent bed track was designed to fit the bed. In short, the bed apparatus insured that a subject could be scanned in all three systems such that repositioning of the subject, relative to the bed, is not necessary. Due to the incompatibility of the bed with the original Genisys system (the dual-headed "G2"), no measurements were taken on this system. All visualization was performed using Amide while all calculations were performed using the Interactive Data Language (IDL) version 7.0.

In order to obtain positioning data within the microCT, a spherical phantom (Ø = 16 mm) was filled with saline and affixed to the bed such that the sphere appeared within the approximate center of the microCT FOV. The sphere was scanned repeatedly over ten acquisitions. Each scan consisted of an 8-minute acquisition using a tube energy of 70 kVp and 500 µA current. The resulting raw images had voxel dimensions of 256 x 256 x 496 (0.2 mm x 0.2 mm x 0.2 mm). In between each acquisition, the bed was fully removed from the scanner and then reinserted. The sphere itself was not repositioned relative to the bed. All centroids were calculated using the same choice of origin. One example of these ten images is shown in Figure 5.5.

Figure 5.5: Example transaxial image of sphere scanned within the microCT scanner. The size of this slice is 256 x 256 (0.2 mm x 0.2 mm). The centroid of this sphere is the quantity of interest, but the nonzero voxels of the bed and nonuniformity within the sphere precludes this. Additional image processing must be performed before the geometrical centroid can be calculated.

The straightforward calculation of the centroid consists of a position average weighted by voxel values (i.e. "center-of-mass"). However, the nonzero voxels of the bed interfere with this. In addition, the nonuniformity within the sphere may bias the results away from the geometric center of the sphere. As a result, all ten images were cropped so that only the sphere was visible. In addition, a voxel threshold was applied so that all voxels within the sphere with a Hounsfield unit value of zero, or greater, were reassigned to have a value of one. (The inclusion of voxels with values less than zero would undesirably include the voxels surrounding the sphere.) All other voxels were set to zero. This simple method insured that the centroid calculation included only voxels within the sphere. The cropped and thresholded image obtained from Figure 5.5 is shown in Figure 5.6.
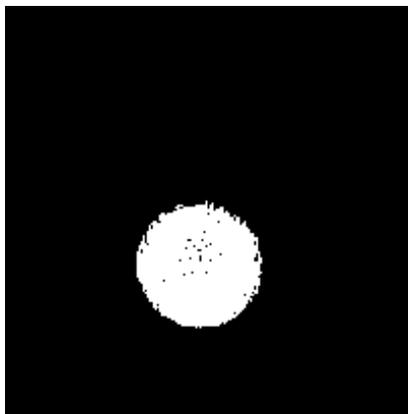
Figure 5.6: Cropping and thresholding applied to the sphere shown in Figure 5.5. All voxels which lie outside the sphere are set to have a value of zero. Most voxels within the sphere have a value of one, but a negligible few a value of zero due to the presence of voxel values below zero in the original image.

As seen in Figure 5.6, most voxels within the original sphere have a value of zero or greater. However, a few have values below zero. As a result, the thresholding is not able to set these voxels to one. In order to compute the centroid, the eight nonzero voxels with the largest and smallest x, y, and z coordinates were found. These coordinates constitute the periphery of sphere. The centroid of the sphere was then computed as the geometric center of these coordinates. In short, this additional processing performed on these images makes computation of the centroid more convenient and less error-prone to the thresholding artifact. Despite these measures, it is still possible that the centroid measurement may introduce errors within the calibration data.

The same sphere ($\varnothing$ = 16 mm) used in the microCT images were also used for measuring position reproducibility of the Inveon microPET system. The sphere was filled with approximately 50 µCi of ^{18}F-Fluorodeoxyglucose (^{18}F-FDG) activity at the beginning of the first scan. Ten images were acquired in a manner similar to that of the microCT; the bed was fully

removed and replaced prior to each acquisition, but the sphere itself remained undisturbed. Each image consisted of a 1-minute acquisition followed by 3D histogramming, and Ordered Subset Expectation Maximization (OSEM) 3D with Maximum A Posteriori (MAP) reconstruction. Decay correction, randoms correction, deadtime correction, and normalization were applied; no attenuation correction was applied. The resulting raw images had voxel dimensions of 128 x 128 x 159 (0.86 mm x 0.86 mm x 0.8 mm). An example of one of these images is shown in Figure 5.7.
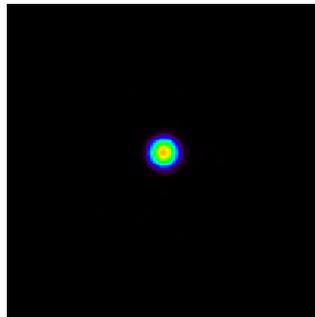
Figure 5.7: Example transaxial image of the spherical phantom scanned with the Inveon preclinical PET scanner. The size of this slice is 128 x 128 (0.86 mm x 0.86 mm). The colorbar has been adjusted to accommodate the entire dynamic range of voxel values. No noticeable artifacts prevent calculation of the centroid.

With the exception of noise, the entire signal within the image originates from the sphere activity itself. Therefore, the straightforward calculation of the centroid (i.e. "center-of-mass) is readily feasible. The last of the ten images was acquired within 30 minutes of the first scan. The decay of activity ($T_{1/2}$ = 109.8 minutes) may cause subsequent images to have nominally different voxel values from previous images. However, due to the symmetry of the sphere, these slight differences should not introduce appreciable error in the centroid calculation. Again, the

115

calculated centroid, due to the symmetrical shape, should coincide with the geometrical center of the sphere.

In addition to testing the positioning precision on the microCT and Inveon, the same spherical phantom (Ø = 16 mm) was used for measurements on the Genisys G4. The sphere was filled with approximately 20 µCi of ^{18}F-FDG. Ten 1-minute acquisitions were implemented. All acquisitions were performed within 15 minutes of each other. As a result, the intermittent ^{18}F decay is considered negligible. Histogramming was performed using decay correction, randoms correction, deadtime correction, and normalization; attenuation correction was not available. Reconstruction was performed using the default EM 3D with 60 iterations. Following reconstruction, images of the Digital Imaging and Communications in Medicine (DICOM) format were produced and converted to raw images using Amide. The final raw images had voxel sizes of 96 x 96 x 208 (0.46 mm x 0.46 mm x 0.46 mm). An example of one of these raw images is shown in Figure 5.8.
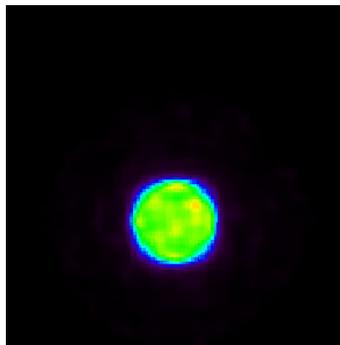
Figure 5.8: Example transaxial image of the spherical phantom scanned with the preclinical Genisys G4 scanner. The size of this slice is 96 x 96 (0.46 mm x 0.46 mm). The colorbar has been adjusted to accommodate the entire dynamic range of voxel values. No noticeable artifacts prevent calculation of the centroid.

After image acquisitions, the 3D centroid of each image was tallied and separated into the three spatial dimensions: left-right, superior-inferior, and anterior-posterior. The results, for each scanner, therefore consisted of 30 individual measurements. The variance between measurements for each of the three dimensions was computed. A total variance for all measurements was also computed. These measurements quantify the position repeatability (i.e. rigidity) of both systems; larger values of variance imply less repeatability, and, therefore, less suitably for the rigid body transformation. These measurements take into account the mechanical repeatability of bed insertion and image statistical noise which may preclude obtaining accurate calibration data, and, therefore, a proper fitting rigid body transformation.

5.2.2 Measurement of Calibration Data

The coregistration algorithm relies on independently acquired calibration data in order to compute a transformation between two different scanners. This calibration data is, in essence, a substitute for the information normally provided by the optical tracker use for motion correction studies. The calibration data consists of coordinates which attempt to measure the same physical point within the coordinate systems inherent in each of these two scanners (Figure 5.3). Initial attempts to measure these points used crudely made "point" sources that were visible on the microCT and microPET systems. However, this method provided inconsistent and error-prone results.

For this reason, we acquired images using the spherical phantom and protocols described in Chapter 5.2.1. The sphere was randomly repositioned relative to the bed after acquisition

within the microCT and microPET systems. In other words, each microCT image has a corresponding microPET image in which the sphere is in the same position relative to the bed. The scanning protocols were identical to those used in the positioning repeatability study. In addition, only the minimum three reference points were used. Centroids of each image were computed identically to that described in Chapter 5.2.1. All three microCT images, fused together, are shown in Figure 5.9. The image shown corresponds to the fusion of images after cropping and thresholding (as seen in Figure 5.6).
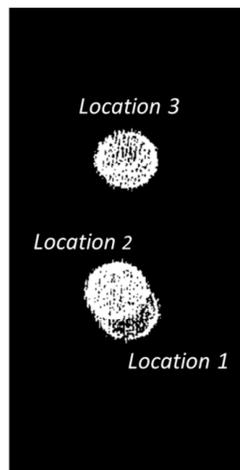
Figure 5.9: Fusion of three separate microCT coronal images of the spherical phantom. The phantom was repositioned in between scans. Centroids of each sphere were computed individually. These images were used to compute the coregistration transformation between the microCT and the Inveon microPET systems.

Images taken at location 1 and location 2 in Figure 5.9 appear to overlap. However, their centroids are distinct and, therefore, provide appropriate calibration data. Note that, due to the thresholding, not all voxels within the sphere are assigned to a value of one. Therefore, the spherical shapes appear "hollow" in certain areas. Although not ideal, this artifact is considered manageable since the geometric center of the sphere is calculated using on the peripheral voxels.

The fused Inveon image corresponding to that in Figure 5.9 is shown in Figure 5.10. (These images have not been coregistered.)
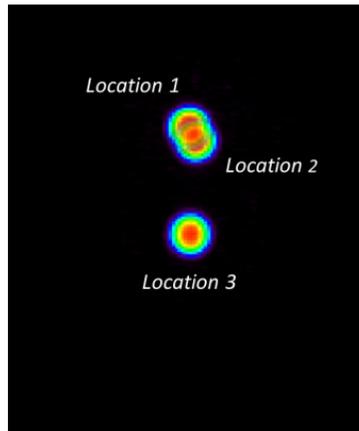
Figure 5.10: Fusion of three separate Inveon coronal images of the spherical phantom. The phantom was repositioned in between scans. The locations indicated correspond to those in Figure 5.9. Centroids of each sphere were computed individually.

Again, overlapping spheres are present, but do not pose a barrier to the accuracy of the calibration data since their centroids do not overlap. Visual comparison of Figure 5.9 and Figure 5.10 indicate that the two images are indeed related by a rotation. The translational component is also present, but not readily seen due to the fact that the choice of origin within each system is arbitrary. In addition, due to the difference in voxel sizes and FOV sizes, a resampling of voxels is also needed. These are all accommodated with the transformation that is computed based on these centroid coordinates.

A nearly identical method was used to acquire calibration data between the microCT and

G4 scanners. The microCT acquisitions were fused and presented in Figure 5.11. The fused

Genisys G4 image, corresponding to Figure 5.11, is shown in Figure 5.12.
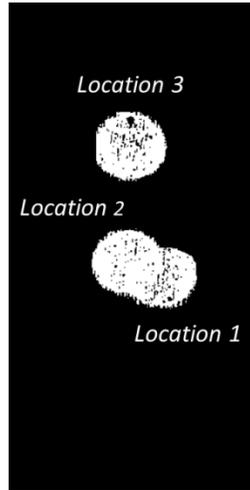
Figure 5.11: Fusion of three separate microCT coronal images of the spherical phantom. The phantom was repositioned in between scans. Centroids of each sphere were computed individually. These images were used to compute the coregistration transformation between the microCT and the Genisys G4 microPET systems.
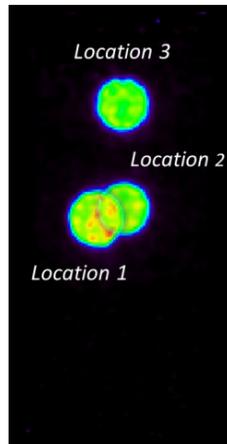
Figure 5.12: Fusion of three separate Genisys G4 coronal images of the spherical phantom. The phantom was repositioned in between scans. The locations indicated correspond to those in Figure 5.11. Centroids of each sphere were computed individually.

Note that the data shown in Figure 5.11 and Figure 5.12 was acquired completely independently of the data shown in Figure 5.9 and Figure 5.10. Therefore, the location labels within Figure 5.9 and Figure 5.10 do not necessarily match those of Figure 5.11 and Figure 5.12. Any resemblance is purely coincidental.

Note that centroids of a sphere are not integral to the computation of the transformation. The geometrical center of the sphere, the centroid of the sphere in the microCT image, and the centroid of the sphere in the microPET images should coincide with one another due to symmetry of the sphere. As a result, computation of the centroid provides a convenient way to identify the same physical point within a set of different imaging systems. The coregistration itself is not reliant on any particular phantom, voxel values, or set of images. In fact, the only input into Horn's method are the coordinates $\vec{r_{Ai}}$ and $\vec{r_{Bi}}$ shown in Figure 5.3. How those coordinates were measured is not a factor in computing the transformation.

5.2.3 Application of the Rigid Body Transformation to Any Pair of Images

The computed transformation results in a rotation, q, and translation $\vec{t}$, as seen in Equation 5.2. Note that this transformation is designed to transform physical coordinates (measured in distance), not voxel values or indices. In addition, both images to be coregistered may not share the same voxel dimensions or FOV size. As a result, the transformation is derived such that the image with the finer voxel size (e.g. microCT) is coregistered to the second image

(e.g. microPET). Referring to Equation 5.1, system A corresponds to the microCT while system B corresponds to the microPET. If performed in an opposite manner, the newly created coregistered image will suffer from undersampling compared to the original (i.e. violate the Nyquist Theorem).

Prior to applying the transformation, a blank image array is created (in IDL) such that its array size matches that of system A (e.g. microCT). This new image array is populated voxel-by-voxel in a raster-like format such that, after completion, the array represents image A coregistered to image B.

Converting each voxel within system A to a coordinate requires the use of the voxels sizes, Δx_A, Δy_A, and Δz_A in each of the three dimensions, respectively. For a given voxel with indices (i_A, j_A, k_A), its corresponding coordinate is computed as

$$\vec{r_A} = (i_A\,\Delta x_A, j_A \Delta y_A, k_A \Delta z_A) \qquad \text{Equation 5.5}$$

Note that, depending on the choice of origin, some components of $\vec{r_A}$ may be negative. Once computed, the corresponding coordinate in system B is found using Equation 5.2. The transformed coordinate, $\vec{r_B}$, must then be converted into voxel indices corresponding to system B. These indices *(i_B, j_B, k_B)* can be computed by simply dividing the components of $\vec{r_B}$ appropriately:

$$(i_B, j_B, k_B) = round\left(\frac{x_B}{\Delta x_b}, \frac{y_B}{\Delta y_b}, \frac{z_B}{\Delta z_b}\right) \qquad \text{Equation 5.6}$$

where x_B, y_B, and z_B represent the three components of $\vec{r_B}$, respectively. The variables Δx_B, Δy_B, and Δz_B represent the voxel sizes (in units of distance) of the images obtained in system B. The *round* operator rounds all values to the nearest integer. This rounding is necessary since fractional indices are computationally meaningless. In short, voxel *(i_B, j_B, k_B)* indicates the voxel in system B which lies the physically closest to voxel (i_A, j_A, k_A) in system A. Once identified, the value of this voxel is simply assigned to the voxel in the previously created empty image. Denoting this new empty image as I_{new}, and the image from system B as I_B

$$I_{new}(i_A, j_A, k_A) = I_B(i_B, j_B, k_B) \qquad \text{Equation 5.7}$$

The process of transforming the coordinate, converting to indices, and assigning voxel values is employed repeatedly for each voxel in system A until all voxels in image A have been accounted for. This new image, I_{new}, has the same orientation, translation, and voxel dimensions as system A by construction; however, it contains information entirely from image B. In short, this new image is a resampled and reoriented version of image B such that it spatially overlaps with image A on a voxel-by-voxel basis. Note the frequency distribution of voxel value is preserved; no values were altered in the coregistration process. In other words, the information (in the mathematical sense) remains essentially constant between image B and the coregistered version of image B. Due to the differences in FOV sizes, it is possible that transformed

coordinates lie outside the measurable FOV; if so, these coordinates are disregarded and no voxel values are assigned to the new image. Finally, note that voxel values from image A are not used anywhere throughout this process. Once completed, this new image is written to a new raw image file. This raw file is then imported into Amide along with image A to produce a fused image.

5.3 Results

5.3.1 Determining Positioning Repeatability

Measurements were taken to quantify the rigidity of the bed apparatus within each scanner. Using the spherical phantom affixed to the bed, ten images were acquired on each scanner. The bed was fully removed and reinserted in between each acquisition. The centroid of each image was calculated and used to represent the position of the bed within the FOV. The variance of all positions, in each of the three spatial directions, was then calculated to quantify how much inherent error the bed apparatus may introduce into the coregistration system. Table 17, Table 18, and Table 19 indicate the computed coordinates and variance for the microCT, Inveon, and Genisys G4 measurements, respectively. Note that the coordinates themselves are arbitrary due to the choice of origin. However, the fact that the results are fairly consistent is of interest.

Measurement Number	Left-Right Direction (mm)	Anterior-Posterior Direction (mm)	Superior-Inferior Direction (mm)
1	23.68	18.67	44.61

	Left-Right Direction	Anterior-Posterior Direction	Superior-Inferior Direction
2	23.68	18.67	44.51
3	23.68	18.67	44.61
4	23.68	18.67	44.61
5	23.78	18.57	44.61
6	23.78	18.57	44.61
7	23.87	18.57	44.61
8	23.87	18.57	44.61
9	23.87	18.57	44.61
10	23.87	18.57	44.61
Variance (mm^2)	8.59×10^{-3}	2.57×10^{-3}	9.66×10^{-4}
Total Variance (mm^2)	1.32×10^{-2}		

Table 17: Bed positioning measurements taken with the Siemens microCT. The results, in general, are quite reproducible.

Measurement Number	Left-Right Direction (mm)	Anterior-Posterior Direction (mm)	Superior-Inferior Direction (mm)
1	54.92	57.20	54.58
2	54.93	57.19	54.54
3	54.91	57.21	54.57
4	54.91	57.20	54.55
5	54.90	57.22	54.52
6	54.91	57.21	54.51
7	54.91	57.22	54.51
8	54.90	57.20	54.50
9	54.89	57.22	54.50
10	54.88	57.21	54.52
Variance (mm^2)	2.29×10^{-4}	8.83×10^{-5}	7.87×10^{-4}
Total Variance (mm^2)	1.10×10^{-3}		

Table 18: Bed positioning measurements taken with the Inveon microPET system. The results, in general, are quite reproducible.

Measurement Number	Left-Right Direction (mm)	Anterior-Posterior Direction (mm)	Superior-Inferior Direction (mm)
1	21.54	20.54	46.26
2	21.54	20.53	46.21
3	21.56	20.54	46.22
4	21.56	20.52	46.22
5	21.57	20.52	46.19
6	21.55	20.53	46.20
7	21.55	20.52	46.16
8	21.55	20.54	46.25
9	21.56	20.54	46.20
10	21.57	20.54	46.19
Variance (mm^2)	1.12×10^{-4}	6.14×10^{-5}	8.63×10^{-4}
Total Variance (mm^2)	1.04×10^{-3}		

Table 19: Bed positioning measurements taken with the Genisys G4 preclinical PET system. The results, in general, are quite reproducible.

5.3.2 Measurement of Calibration Data

Three separate images of the spherical phantom were acquired for each calibration. For the microCT-Inveon calibration, the phantom was randomly repositioned in between each pair of microCT and Inveon scans. Figure 5.9 and Figure 5.10 represent the fusion of all three images taken within each system. Centroids of each image were calculated separately. These centroids were then used as coordinates for computing a rigid body transformation via Horn's method as described in Equation 5.2 and Figure 5.3. The resulting χ^2 was computed as 1.5 x 10^{-3} mm^2. The resulting transformation included approximate values of $q \approx$ (0.0, 1.0, 0.0, 0.0) and t $\approx$ (27.2, 75.9, 99.7) mm. The transformation represents the conversion of coordinates within the microCT frame to the microPET frame (as opposed to the microPET frame to the microCT frame). The transformation was separately applied to these. For visual clarity, these coregistered images are shown separately. Figure 5.13, Figure 5.14, and Figure 5.15 correspond to location 1, location 2, and location 3, respectively. The location numbers correspond to those seen in Figure 5.9 and Figure 5.10.
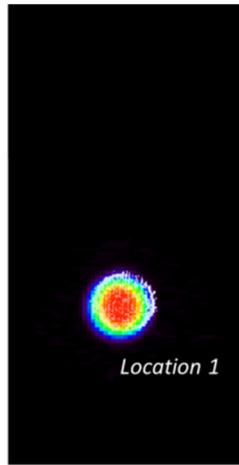
Figure 5.13: Coregistered coronal images of spherical phantom within the microCT and Inveon at location 1. The sphere at location 1 in Figure 5.10 has been transformed to coincide with the sphere at location 1 in Figure 5.9.

Figure 5.14: Coregistered coronal images of spherical phantom within the microCT and Inveon at location 2. The sphere at location 2 in Figure 5.10 has been transformed to coincide with the sphere at location 2 in Figure 5.9.
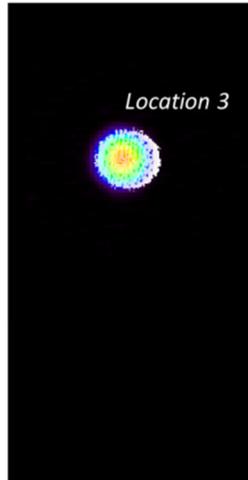
Figure 5.15: Coregistered coronal images of spherical phantom within the microCT and Inveon at location 3. The sphere at location 3 in Figure 5.10 has been transformed to coincide with the sphere at location 3 in Figure 5.9.

In addition to coregistration of the Inveon images to the microCT images, coregistration was also performed between the microCT and Genisys G4 systems in a very similar manner. The resulting χ^2 was computed as 0.31 mm^2. The resulting transformation included approximate values of $q \approx (0.0, 0.0, 0.0, 1.0)$ and $t \approx (45.2, 42.6, -7.9)$ mm. For visual clarity, each individual coregistration is shown separately. The results are shown in Figure 5.16, Figure 5.17, and Figure 5.18. The location numbers correspond to those seen in Figure 5.11 and Figure 5.12.
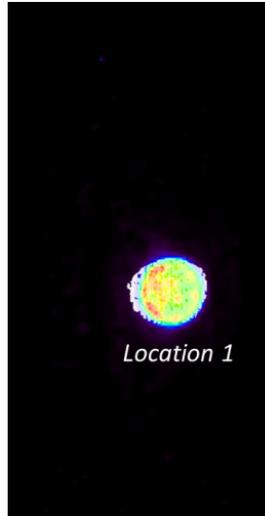
Figure 5.16: Coregistered coronal images of spherical phantom within the microCT and Genisys G4 at location 1. The sphere at location 1 in Figure 5.12 has been transformed to coincide with the sphere at location 1 in Figure 5.11.
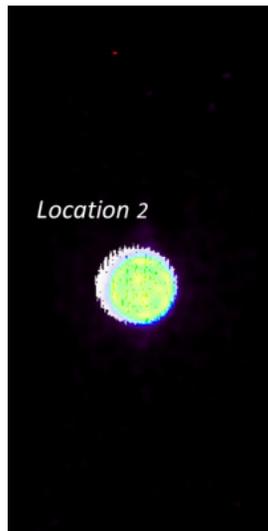
Figure 5.17: Coregistered coronal images of spherical phantom within the microCT and Genisys G4 at location 2. The sphere at location 2 in Figure 5.12 has been transformed to coincide with the sphere at location 2 in Figure 5.11.
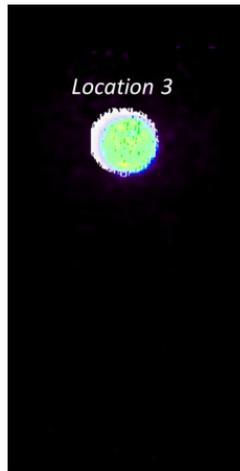
Figure 5.18: Coregistered coronal images of spherical phantom within the microCT and Genisys G4 at location 3. The sphere at location 3 in Figure 5.12 has been transformed to coincide with the sphere at location 3 in Figure 5.11.

Coregistration of these spherical phantom images is not necessary; once computed, the transformation obtained from Horn's method can be used to coregister any pair of images taken from the same systems using the same bed positioning and FOVs. However, coregistration of these images provides a visual qualitative check for consistency and accuracy. Any errors present within this calibration data will also be visible when coregistering images of interest. There is an overall agreement between locations of each sphere after coregistration.

5.3.3 Application of the Rigid Body Transformation to Any Pair of Images

Using the transformation derived with the spherical phantom images and Horn's method, pairs of mice images were coregistered. The scanning protocols were similar to those used for

the calibration data. However, the microPET scan used approximately 70 µCi of ^{18}F-FDG; the 1-minute scan started approximately one hour post-injection. The microCT scan, Inveon scan, and coregistered images can be seen in Figure 5.19, Figure 5.20, and Figure 5.21, respectively.

Figure 5.19: Raw CT image slices of a normal C57BL6 mouse, scanned on the Siemens MicroCT II system, using Amide. The voxel sizes are 256 x 256 x 496 (0.2 mm x 0.2 mm x 0.2 mm).
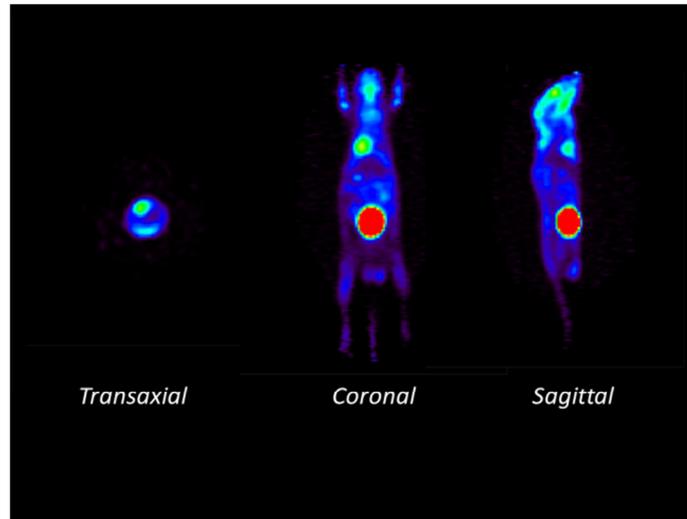
Figure 5.20: Raw PET image slices of a normal C57BL6 mouse, scanned on the Siemens Inveon microPET system, using Amide. The voxel sizes are 128 x 128 x 159 (0.86 mm x 0.86 mm x 0.8 mm). This scan was performed immediately after the scan shown in Figure 5.19 without repositioning the mouse relative to the bed.
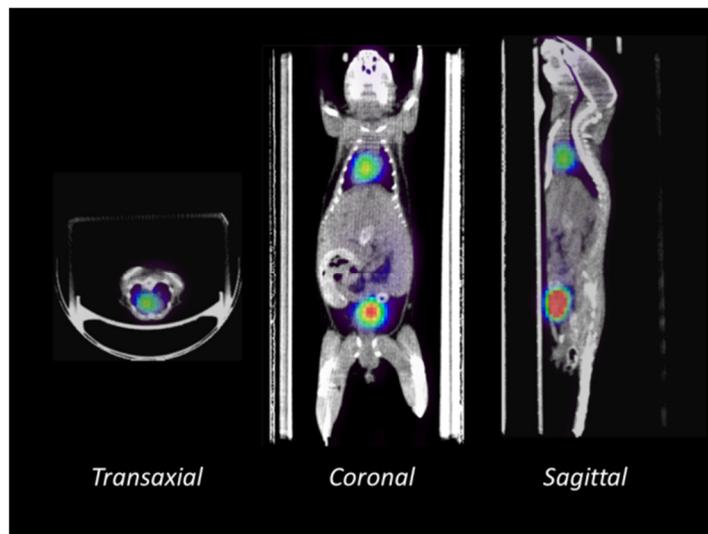
Figure 5.21: Coregistration between microCT and Inveon images. The Inveon microPET image (Figure 5.20) was transformed to match the orientation and sampling of the Siemens microCT II image (Figure 5.21). After coregistration, the fused image was reoriented to match that of the standard orientation.

Coregistration of a mouse between the microCT and the Genisys G4 was performed on another mouse using the same method. Calibration data, however, was acquired completely independently of that of the previous study. This study consisted of a NOD SCID Gamma (NSG) mouse injected with a CEM-shC tumor cells on its left, and CEM-shdCK tumor cells on its right. The mouse was intraperitoneally injected with approximately 200 μL of Captisol about four hours prior to scanning. The microCT scan is shown in Figure 5.22. Without repositioning the mouse relative to the bed, the mouse was also scanned on the Genisys G4. Approximately 26 μCi of ^{18}F-2'-deoxy-2'-arabinofuranosyl-cytosine (D-FAC) was injected about one hour prior to the ten minute scan. The resulting raw image is shown in Figure 5.23. The Genisys G4 image in Figure 5.23 was coregistered to the microCT image in Figure 5.22. This coregistered image was then put into the standard orientation. The fused image is shown in Figure 5.24.
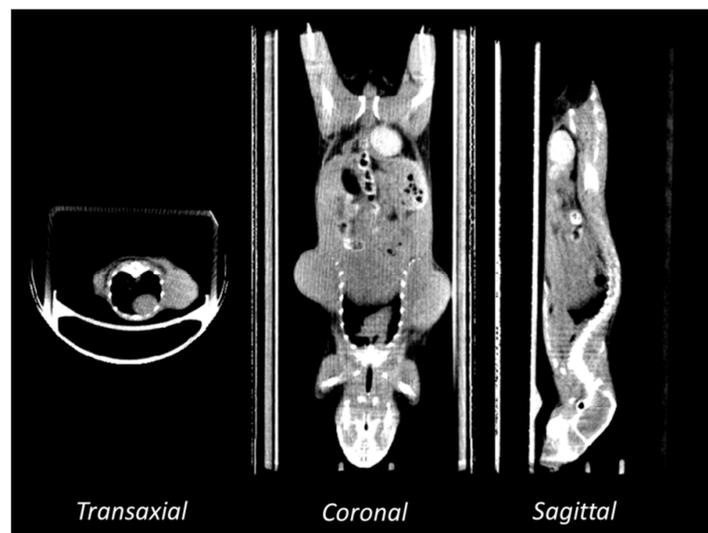
Figure 5.22: Raw CT image slices of a NSG mouse with two laterally-positioned tumors, scanned on the Siemens MicroCT II system, using Amide. The voxel sizes are 256 x 256 x 496 (0.2 mm x 0.2 mm x 0.2 mm).
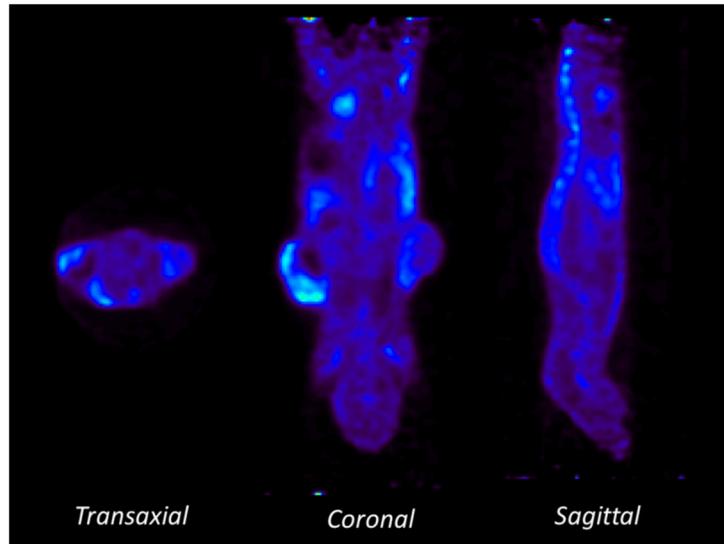
Figure 5.23: Raw PET image slices of a NSG mouse, scanned on the Genisys G4 microPET system, displayed using Amide. The voxel sizes are 96 x 96 x 208 (0.46 mm x 0.46 mm x 0.46 mm). This scan was performed immediately after the scan shown in Figure 5.22 without repositioning the mouse relative to the bed.
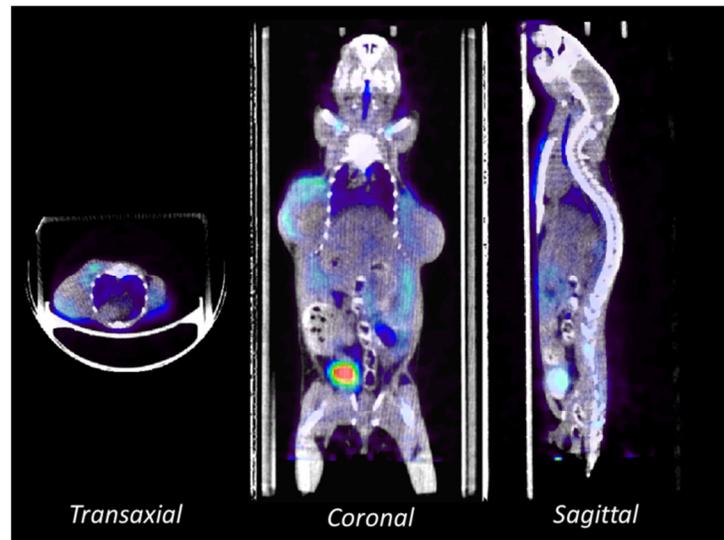
Figure 5.24: Coregistration between microCT and Genisys images. The Genisys G4 microPET image (Figure 5.23) was transformed to match the orientation and sampling of the Siemens microCT II image (Figure 5.22). After coregistration, the fused image was reoriented to match that of the standard orientation. The artifacts visible near the hind legs are not due to coregistration, as can be seen in Figure 5.23.

5.4 Discussion

5.4.1 Positioning Repeatability

Table 17 through Table 19 list the positioning results for the spherical phantom when scanned removed and reinserted into each system ten times. All three tables indicate a high degree of repeatability. In fact, the total standard deviation within each system is considerably lower than the spatial resolution of the system. The microCT bed apparatus appears to have the greatest variance. It is not clear, however, if this is due to undesired latitude in the bed apparatus itself, or due to any error introduced by the cropping and thresholding processing performed on the spherical phantom images. As seen in Figure 5.6 and Figure 5.9, the thresholding does not properly separate all voxels within each phantom image from their empty surroundings. As a result, computing the center of mass may be somewhat error prone. Note that the microCT positioning (Table 17) was not only reproducible, but chronologically consistent. For example, the last four measurements were identical within $1/100^{th}$ of a mm. This is not seen in the Inveon positioning (Table 18) or Genisys G4 positioning (Table 19). In fact, none of these measurements were exactly reproducible in all three dimensions. This may seem a little unexpected given the fact that no processing of the microPET images was performed prior to computing the centroid. This may be explained by the generally poor signal statistics of PET compared to that of CT.

For all of these repositioning studies, the bed actuators built into each scanner were not used; the bed itself was slipped in manually until firmly docked with the locking mechanism. Had the actuators been used for each scan, variability might have been markedly increased due to

135

the limited precision with which the bed can be positioned. In short, the method of bed positioning seems very favorable for rigid body coregistration purposes; only minimal error should be introduced by the bed apparatus itself. Assuming statistical independence of the positioning error within systems, the total expected error may be approximated by the sum of the variances.

Future improvements to this method may consist of imaging using a CT contrast agent within the sphere or designing a new phantom to accommodate this method (forgoing the spherical phantom altogether). In addition, the spherical phantom was placed in the approximate center of each FOV. It is possible that repositioning may exhibit more variability for positions near the edge of the FOV due to imprecision within the bed tracks. Future studies may confirm or disprove this by repeated measurements similar to those performed. Positioning measurements may also be conducted without relying on the images at all using external systems similar to those used in PET brain motion correction.

5.4.2 Measurement of Calibration Data

The calibration data obtained from the separate sphere centroids were fed into Horn's method in order to produce a transformation. Two separate transformations were computed for the microCT-Inveon coregistration and microCT-Genisys coregistration. Each transformation was based on independently acquired calibration data.

The quaternion calculated for microCT-Inveon coregistration, $q \approx (0.0,\ 1.0,\ 0.0,\ 0.0)$, indicates a 180^0 rotation about the first axis (i.e. left-right axis) of the microCT image. This can

be visually verified using Figure 5.9 and Figure 5.10. In this particular example, a rotation of 180^0 about a given axis was found; the coregistration method, in general, however is not restrained to such angles. Horn's method will compute any arbitrary rotation about any arbitrary axis with equal accuracy due to the fact that it uses physical coordinates. Although most rotations likely to be encountered between imaging systems are similar to those computed in this example, the rotations are also influenced by any mispositioning of the bed tracks within the bore. For example, if the bed tracks do not run exactly parallel to the axial direction of the bore, the computed rotation may not yield such results even if no other sources of error are present.

The transformed spherical images are shown in Figure 5.13 through Figure 5.15. All images indicate that the Inveon images are properly oriented and shifted such that the spheres of the microCT images and microPET images generally coincide as desired. The resampling of the microPET spheres appears to be reasonable given the fact that the sizes of the coregistered spheres and the microCT spheres are approximately equal. However, there is a slight discrepancy visible on all three figures. In each case, it appears as if though the computation of translation slightly underestimated the shift in the left-right direction. This may be due to inaccuracies of the centroid measurements (Figure 5.4). However, it may also be due to the roundoff error associated with the discretized voxels (Equation 5.6). In short, if the voxel sizes of system A are not integer multiples of voxel sizes from system B, this roundoff error may be present. This is true for voxel sizes of the microCT and microPET systems; voxel sizes of the microCT are 0.2 mm x 0.2 mm x 0.2 mm while voxel sizes from the Inveon system is 0.86 mm x 0.86 mm x 0.8 mm. Due to this roundoff error, the order in which system A coordinates are "looped through" (Equation 5.5) may have a minor effect on the final result.

The results of the microCT-Genisys coregistration follow a similar pattern. The quaternion computed for the microCT-Genisys coregistration, $q \approx (0.0, 0.0, 0.0, 1.0)$, indicates a 180^0 rotation about the third axis (i.e. superior-inferior axis) of the microCT image. A visual comparison between Figure 5.11 and Figure 5.12 confirms this. When coregistering the individual spherical images, a slight error is visible. Figure 5.16, Figure 5.17, and Figure 5.18 all indicate a general agreement between spherical images, but also show a tendency for the Genisys G4 image to be shifted slightly in the left-right direction relative to the microCT image. The potential reasons for this shift are identical to those of the microCT-Inveon data set; errors within the centroid measurements or voxel roundoff errors may both be the source of this discrepancy.

The same direction of erroneous shift can also be seen on the microCT-Inveon coregistered images (Figure 5.13, Figure 5.14, and Figure 5.15). However, note that the former case, the Genisys G4 spheres are shifted to the right of the microCT; in the latter case, the Inveon spheres are shifted to the left of the microCT spheres. In other words, both sets of data include similar, but uncorrelated, error.

Future work to mitigate this may consist of oversampling both images such that this voxel size condition is satisfied. In short, it is clear that the resulting transformation performs the generally desired result. Incremental improvements may be obtained with more precise measurements from the phantom as well as from finer sampling of the images.

5.4.3 Coregistration Applied to Mice Studies

The computed transformation was applied to a pair of mice images taken independently of the calibration data. Coregistration of an Inveon microPET scan to a microCT scan is shown in Figure 5.21. As expected from the calibration data (Figure 5.13, Figure 5.14, and Figure 5.15), the fused image shows strong agreement between the heart and bladder. With adjustments made to the microPET color mapping (not shown), a discrepancy similar to that in Figure 5.13 through Figure 5.15 can be seen. This consistency is simply due to the rigid nature of the computed transform. In other words the magnitude, and general direction of an error introduced when computing the transformation is not amplified, or suppressed, when applying the transformation. The coregistered microPET image may also appear pixelated; this is due to the fact that nearest neighbor interpolation was used (Equation 5.6) in the oversampling. This effect can be easily alleviated using a smoothing filter; for illustration purposes, it has been left unprocessed.

The coregistration between the microCT and Genisys G4 systems also show a generally strong overlap as seen in Figure 5.24. However, as in the case of the calibration data (Figure 5.16 through Figure 5.18), a discrepancy is visible particularly with the bladder in the sagittal view (Figure 5.24). Note that the shift of the G4 image relative to the microCT is similar to that seen in the calibration data. As in the case of the microCT-Inveon coregistration, this may be mitigated using more accurate centroid measurements for the calibration. In addition, the pixilation visible in the Inveon image within Figure 5.21 is not as pronounced within the G4 image of Figure 5.24. This is not specific to the coregistration; it is simply due to the fact that the voxel sizes of the G4 are closer to the microCT than that of the Inveon. In other words, finer sampling of the G4 image is possible when coregistering to the microCT compared to that of the Inveon.

5.5　Summary

Using methods adapted from PET brain motion correction, a rigid body coregistration method has been developed for small animal imaging. However, no tracking system is available for coregistration purposes. Instead, a calibration method has been developed. The method can be applied to any pair of modalities taken from any pair of systems, even with varying orientation, voxel sizes, and voxel dimensions. The validity of the algorithm depends almost entirely on the assumption of little to no motion of the subject relative to the bed in between acquisitions. At least three randomly chosen points on the bed are selected; coordinates of these points are measured in both systems using the centroid of a spherical phantom. These coordinates are fed into Horn's method which provides the actual transformation to be used for coregistration. The transformation itself consists of rotation and translation components with which one image can be directly fused to another. Once computed, the transformation is applied on a voxel-by-voxel basis to one of the images in order to identify the corresponding voxel within the other image. The values of these corresponding voxels are systematically collected into a new 3D image. This new image can then be directly fused over the existing image. By construction, this new image has the same orientation, voxel size, and voxel dimensions as one image, but contains voxel values derived directly from the other image.

Measurements have shown that the bed apparatus used introduces negligible error into the coregistration; bed positioning within each FOV is quite reproducible. The main source of error is believed to be the method by which calibration data is obtained; the centroids of the spherical phantoms may not provide the most reliable measurements of coordinates within two different systems. Voxel roundoff error is thought to contribute to the error as well. The

transformation derived from the calibration data appears to be in strong agreement with the positioning of the randomly sampled points. When applied to the mice images, this same transformation provides generally positive results. The minor visible discrepancy within the calibration data is similar to that seen within the mice images. This consistency is expected due to the nature of rigid body transforms; any errors present within the calibration will be apparent in the same manner during coregistration.

Future research into this method will focus on obtaining more accurate coordinate measurements for the calibration. This can be performed by using smaller spheres, or by constructing customized phantoms which would provide localization for all modalities of interest. If the computed coordinates are believed to have some error, an iterative method for refining them may also be derived due to the convexity of the least squares criterion [14]. Finally, fusion of three or more images may also be obtained using the methods described or by using appropriate combinations of transforms and inverse transforms.

6   Spatial Linearity on the Genisys System

6.1   Introduction

6.1.1 Quantifying Spatial Linearity

By definition of a rigid transformation, the Euclidean distance between any two points is identical before and after the transformation. Comparison of the distance between physical points within the subject and the corresponding points within the image may reveal this not to be true. For example, points near the edge of the field-of-view (FOV) may be distorted due to the system configuration. As a result, these distances may be altered; the relative position of a given point within the subject may not be properly positioned within its image. In short, a given point within the subject (relative to some origin) and within the image (relative to the same origin) may not be identical [26]. Mathematically, the process of producing an image from a given physical subject may introduce non-rigid features such as shearing, dilation, warping, etc.; in essence, any transformation other than a pure rotation and translation is considered non-rigid.

Measuring this spatial linearity can be performed by imaging a grid of equally spaced points, and then comparing the position of each point on the grid to its corresponding point in the image. The regularly spaced physical grid and its randomly distorted image are graphically shown in Figure 6.1.
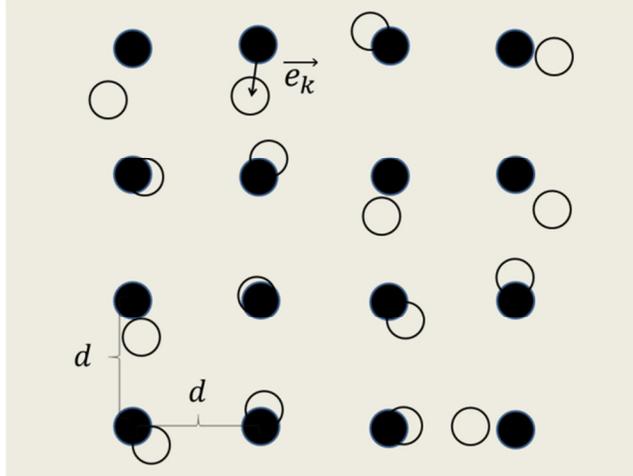
Figure 6.1: Graphical illustration of measuring the spatial linearity of an imaging system. Positions of the regularly-spaced physical points are shown in solid black; their corresponding measured positions within the image are shown as empty circles. The error between each grid point and its corresponding image is measured as a vector $\vec{e_k}$. All adjacent grid points are spaced by a fixed distance d. Although the concept is applicable in 3D, it is shown in 2D for visual clarity.

Within Figure 6.1, each point of the grid (solid black circle) has a corresponding point within the image (empty circle). Although shown in two dimensions (2D) for visual clarity, the concept can be extended into three dimensions (3D). Note that, in general, they do not overlap due to the non-rigid distortions introduced by the imaging process. The distance between adjacent grid points (i.e. grid spacing) is labeled as d. The error between the position of the imaged point and its expected position based on the grid points is labeled as a vector $\vec{e_k}$. This error measures the distance between the center of the solid circle and the center of its corresponding empty circle. For the sake of generality, the distortions shown are random. In practical applications, some specific forms of distortions, such as pincushion distortion, may be very apparent. Regardless, the level of distortion may be quantified in the same manner. By adding the errors in quadrature, and taking the average over all n points, we obtain a total mean-squared error, χ^2.

$$\chi^2 = \frac{\sum_{k=1}^{n}|\vec{e_k}|^2}{n}$$

<div align="right">Equation 6.1</div>

The metric shown in Equation 6.1 can be interpreted as the total error, or variance, in all three dimensions. It has units of distance squared (i.e. mm^2). The quantity in absolute value is the Euclidean norm of the error for a given pair of points. The addition is performed in quadrature which implies that the position error for each imaged point is independent of all other imaged points.

6.1.2 Using Rigid Body Coregistration to Measure Spatial Linearity

A direct approach to measuring this error (Equation 6.1) would require the locations of each grid point. This location, of course, must be measured independently of the actual image. Note, however, that Figure 6.1 can be interpreted as a coregistration between two images. The first image is a synthetic grid with adjacent points spaced exactly by d. The second image is the measured image of this synthetic grid. Both images are shown side-by-side in Figure 6.2.
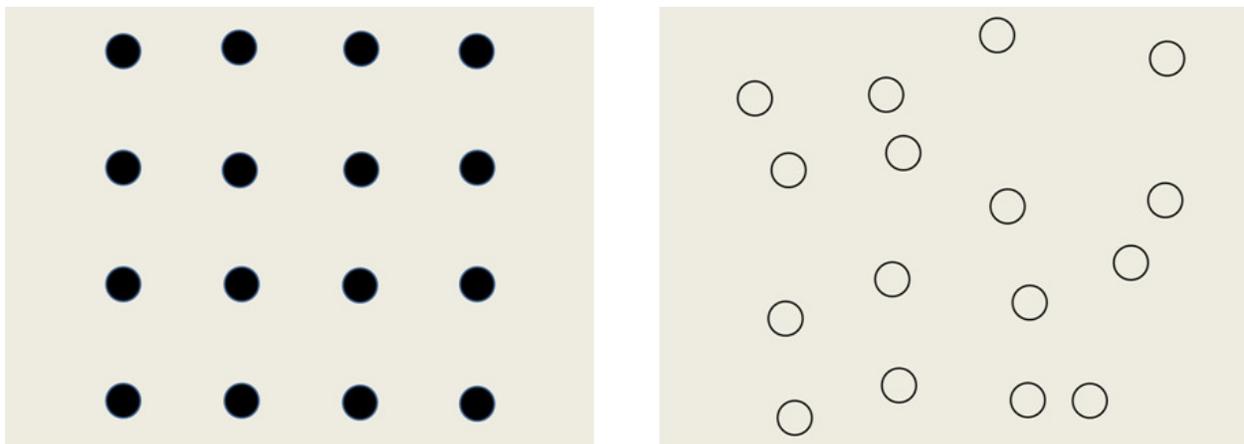
Figure 6.2: The synthetic grid composed of points with consistent grid spacing d is shown on the left. The image obtained from measured positions of these dots in shown on the right. The spatial linearity measurement shown in Figure 6.1 can be obtained by a rigid coregistration between this synthetic grid and the measured image.

Note that the coregistration method (Chapter 5.1.5) performs this rigid coregistration using Equation 6.1 as the criterion for best fit. In other words, using Horn's method [23, 24] to perform a coregistration between this synthetic image and a measured image of grid points minimizes the least-squares metric in Equation 6.1. The least-squares metric used in the coregistration method (Equation 5.3) relies on randomly selected points within the FOVs for two systems. It should only be interpreted as a relative measure of how well a transformation will coregister two images. The metric shown in Equation 6.1, however, can be interpreted as an absolute measure of how well the synthetic grid and imaged grid are coregistered; no random points are used for these purposes. Note that the particular coordinate system used in localizing points is irrelevant; Horn's method will inherently apply a rotation and translation such that the agreement between the synthetic points and the measured points will be maximized in the least-squares sense.

Horn's method only makes use of coordinates, as opposed to images. Therefore, no synthetic image is actually required; only the coordinates (relative to any arbitrary point) of the regularly-spaced points are necessary. Coordinates of the points within the measured imaged are defined by their centroids. These coordinates may also be relative to any arbitrary point. If discrepancies between coordinate systems are present, they will be automatically accommodated for within Horn's method via a rotation and translation. Unlike the metric produced within the calibration data for coregistration, this metric is absolute. In other words, higher values of χ^2 necessarily imply less spatial linearity.

Due to the observed relationship between Figure 6.1 and Figure 6.2, a method of creating a regularly-spaced grid and measuring the centroids of each imaged point within the grid was developed in order to assess spatial linearity of the Genisys dual-headed ("G2") system (Figure 2.1). Regularly-spaced coordinates within the synthetic grid were computationally generated. A second set of corresponding coordinates of the centroids within the measured image were then computed. Both sets of coordinates were used as input into Horn's method in order to compute the mean-squared error metric (Equation 6.1). Although a transformation is also computed, it is not used for purposes of measuring spatial linearity.

6.2   Method

6.2.1 Printing of Grids

In order to compute the spatial linearity as described, a grid of regularly spaced activity "dots" is needed for scanning on the Genisys system. Due to the difficulties in producing a 3D grid, initial attempts were restricted to simple 2D grids. These 2D grids were created by printing activity onto high-quality photographic paper using a standard inkjet printer. The graphical software Canvas (ACD Systems, Victoria, BC, Canada) version 10 was used for all printing needs. Using a syringe, approximately 50 µCi of ^{18}F-Fluorodeoxyglucose (^{18}F-FDG) was injected into the blank ink cartridge. Soon after, grids of various sizes were printed using only black ink. After printing, grids were laminated for additional rigidity. Potential sources of error involved in this process stem from the fact that it is unclear how much ink within the cartridge is

used in the actual printing. It is also unclear how consistent the activity of each printed dot is. This quantity, however, may be estimated from the final image.

For sufficient counts within the final image, larger diameter dots are desirable. However, for computation of the centroid within these dots of unknown activity distribution, smaller dots are desirable. In addition, dots smaller than the spatial resolution of the system would be unreliably imaged. In addition to the size of dots, the spacing between dots must also be considered. Smaller spacing allows for more dots which, in turn, allow for a more comprehensive measure of spatial linearity. In addition, Horn's method has a minimal requirement of three points to be used, but can accommodate a virtually limitless number of points. While the execution time may increase slightly for a higher number of points, the accuracy and computational method do not change. Two slightly different grids were produced. The first grid consisted of 260 dots (Ø = 1.0 mm) arranged in 13 columns x 20 rows. The grid spacing (i.e. *d* in Figure 6.1) was 5.0 mm. A photograph of this printed grid is shown in Figure 6.3.
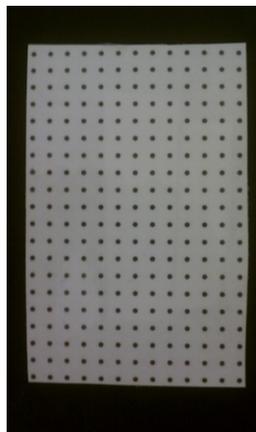
Figure 6.3: Photograph of the printed grid consisting of 260 dots. Each dot is 1.0 mm in diameter, and spaced 5.0 mm away from adjacent dots. The grid was laminated in order to increase rigidity. After imaging of the grid, the centroid locations of each dot within in the FOV

were computed. The agreement between these positions and the hypothetical positions, based on the 5.0 mm separation, was then quantified using Horn's method.

The second grid consisted of 200 dots ($\emptyset$ = 2.0 mm) arranged in 10 columns and 20 rows. The grid spacing for this second set was 6.0 mm. No visual differences, in terms of ink uniformity or contrast, were apparent between dots on either grid.

6.2.2 Imaging Protocols for Grids

Both grids were scanned on the Genisys in an identical manner. In order to reduce position range (and, therefore, annihilation within the air surrounding the grid), imaging was performed by placing the grid sandwiched between two plates of metal. The first grid was sandwiched with plates of Galvanized steel with 0.5 mm thickness; the second grid was sandwiched in between plates of aluminum (Al) with 1.0 mm thickness each. Each plate had a length and width similar to the grids. The grids were placed in between the plates such that the plates covered both faces. In addition to reducing positron range, the use of plates also served to insure that the grid remained flat during scanning; any curvature of the grid would effectively distort the regular spacing and introduce errors within the spatial linearity.

While the positron range from ^{18}F decay (maximum energy of 634 keV) is well known to be ~2 mm in water, positron ranges in Galvanized steel and aluminum were estimated using data published by the National Institute of Standard and Technology (NIST) Physical Measurement Laboratory [27]. Using iron as a substitute for steel and approximating the positron emission

energy as 600 keV, the positron projected range for steel was recorded as 2.82 x 10^{-3} g/cm^2 while the positron projected range for Al was recorded as 2.07 x 10^{-3} g/cm^2. The measured densities of the steel and Al plates were found to be 7.5 g/cm^3 and 2.7 g/cm^3, respectively. Using these measured densities and the NIST data, the positron range within steel and Al can be estimated to be 3.8 x 10^{-3} mm and 7.6 x 10^{-3} mm, respectively. Since the thickness of all plates is at least two orders of magnitude greater than these estimated ranges, positron annihilation within the plates is nearly guaranteed.

The plates and grid were placed in the approximate center coronal slice of the system in order to minimize the effects of limited tomography. Rather than using the conventional bed, the grid and plates were suspended from a hollow carrier designed to insert flat rectangular sources into the Genisys. Both sides of the plates had minimal contact with the carrier; as a result, photon attenuation between the edge of the plates and the detectors by any surrounding material was virtually eliminated. Acquisition was performed over an eight hour period in order to obtain as many counts as possible. Histogramming included the default 3D method as well as normalization, decay correction, randoms correction, and deadtime correction. No attenuation correction was applied. Although attenuation might be introduced by the plates, it is believed that this effect would be minimal compared to the loss of resolution from positron annihilation within the surrounding air. Reconstruction was performed using the default Expectation Maximization (EM) algorithm with 100 iterations and a post reconstruction smoothing filter (σ = 0.35 mm). The final raw images produced from the Genisys had voxel sizes of 80 x 80 x 176 (0.55 mm x 0.55 mm x 0.55 mm). Note, however, that only a few of the slices near the central coronal slice had appreciable activity due to the placement of the grid within the scanner.

6.2.3 Computation of the Dots' Centroids and Total Mean Squared Metric

Following imaging of each grid, the image was viewed within Amide. Due to the large size of the grids relative to the size of the FOV, some dots fell out of the FOV and were not imaged. Dots imaged near the edge of the FOV exhibited artifacts; the voxel indices of such dots were noted and restricted from inclusion in the centroid analysis. Computation of the remaining dots, towards the center of the FOV, was performed in a systematic raster-like format using the Interactive Data Programming Language (IDL) version 7.0. Further analysis was also restrained to the coronal slice which exhibited the highest count level; all other coronal slices were ignored. In essence, the 3D image volume was reduced to a 2D slice based on total voxel values within each coronal slice. Inspection of these images has shown that the vast majority of counts within the 3D volume were included within the selected 2D slice.

For each dot, a square region of interest (ROI), with dimensions d x d, was defined around each dot such that the dot lay in the approximate center of the ROI. Separate ROIs were defined for each dot such that the ROI "tiles" were disjoint and contained only one dot. The centroid of each dot was computed as the average of each voxel position, within the ROI, weighted by voxel value (i.e. "center-of-mass"). In other words, each voxel within the 2D slice contributed to the centroid computation of only one dot. In addition, the minimum and maximum errors contributed from each dot (i.e. the minimum and maximum values of $|\vec{e_k}|^2$ in Equation 6.1) were also recorded in order to give an indication of the variability of mispositioning.

After the centroids of all relevant dots were computed, a corresponding set of synthetic coordinates with spacing d was created. Although arbitrary, the initial coordinate generated was

(0.0 mm, 0.0 mm); subsequent coordinates were chosen by adding d to either or both components. The order in which these synthetic coordinates were generated matched that of the raster-like ROI analysis of the centroid computation. Once the number of synthetic coordinates matched that of the number of centroids computed, both sets were input into Horn's algorithm to determine spatial linearity. Horn's algorithm was used such that the measured coordinates of the imaged dots were coregistered to the synthetic coordinates.

6.3   Results

Raw images, produced by the Genisys, were imported into Amide. The 2D slices selected for analysis are shown for the first and second grid in Figure 6.4 and Figure 6.5, respectively.
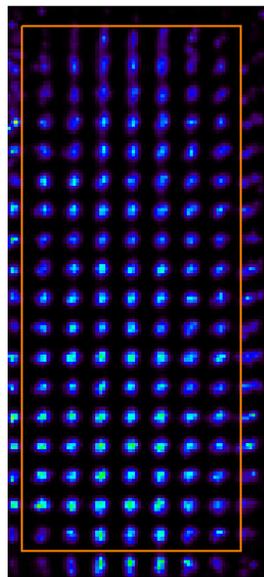
Figure 6.4: Coronal slice used in spatial linearity measurement taken from the Genisys image of the first grid using Ø = 1.0 mm dots with 5.0 mm grid spacing. The slice is 80 x 176 voxels (0.55 mm x 0.55 mm). A photograph of the printed grid is shown in Figure 6.3. The dots near the edge of the FOV exhibit significant distortion and, therefore, were excluded from the centroid analysis. The 126 dots within the rectangular ROI were used in the analysis.
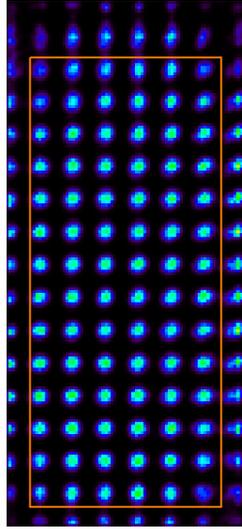
Figure 6.5: Coronal slice used in spatial linearity measurement taken from the Genisys image of the second grid using Ø = 2.0 mm dots with 6.0 mm grid spacing. The slice is 80 x 176 voxels (0.55 mm x 0.55 mm). The dots near the edge of the FOV exhibit significant distortion and, therefore, were excluded from the centroid analysis. The 84 dots within the rectangular ROI were used in the analysis.

The results of the coregistration were obtained for each of the two grids in an independent

manner. The results are shown in Table 20.

Dot Ø (mm)	# of Dots in Printed Grid	# of Dots Visible in FOV	# of Dots Used in Analysis	Grid Spacing (d) (mm)	Plate Material	Plate Thickness (mm)	Total Mean Squared Error (χ^2) (mm^2)	Minimum Error Min$\{|\vec{e_k}|^2\}$ (mm^2)	Maximum Error Max$\{|\vec{e_k}|^2\}$ (mm^2)
1.0	260	171	126	5.0	Galvanized Steel	0.5	0.17	9.8 x 10^{-10}	2.32
2.0	200	136	84	6.0	Aluminum	1.0	0.06	1.4 x 10^{-7}	0.43

Table 20: Results from the spatial linearity measurements. For both grids, the average deviation between the imaged dot and synthetic grid point (i.e. the square root of χ^2) was less than the spatial resolution in the coronal plane (~1.5 mm).

6.4   Discussion

The imaged grids in Figure 6.4 and Figure 6.5 do not show any appreciable distortions near the center of FOV. The dot-to-dot spacing appears to be consistent. Note that the rotation of the grid relative to the FOV is not of consequence when computing the spatial linearity; Horn's method will implicitly rotate and translate the image such that the final result is the best fit between the grid image and synthetic coordinates. The total mean squared error is the metric for this fit.

Table 20 indicates that the spatial linearity, as measured by the total mean squared error, is well within the coronal spatial resolution of the Genisys (~1.5 mm at FWHM) for both imaged grids. The average deviation of each image point from its grid point is approximated by taking the square root of χ^2. This results in average deviation values of 0.41 mm and 0.24 mm for the first and second grid, respectively.

Somewhat surprisingly, the total mean squared error for the first grid is about 4 times larger than that of the second grid even though the dot diameter for the first grid (1.0 mm) is smaller than that of the second grid (2.0 mm). In other words, the synthetic coordinates did a poorer job of predicting where the centroids would be measured in the first grid compared to the second grid. This may be attributed to the fact that more dots were used in the analysis for the first grid (126 dots vs. 84 dots). These additional points were in the periphery of the FOV and, therefore, more prone to distortion. In addition, the maximum error for the first grid is significantly higher than that of the second grid. In fact, this maximum error (2.32 mm^2) implies a deviation of 1.52 mm which is comparable to the coronal spatial resolution. Deviations of this magnitude, if present in significant portions within the FOV, should be considered intolerable for imaging purposes.

6.5 Summary

The spatial linearity for the Genisys system appears to be quite consistent within the central coronal FOV. Discrepancies, as measured by the total mean squared error, were found when using slightly different configurations for the printed grids and metal plates. However, in either case, the distortions present are significantly small compared to the coronal spatial resolution of the system.

Future work should consist of measurements within coronal planes away from the central region where spatial linearity is expected to degrade. A more comprehensive investigation may make use of a 3D phantom grid in order to accommodate the entire FOV as well as to bypass the uncertainties associated with printing the grids. If constructed grids can imaged on other modalities, this method of measured spatial linearity may be applied to them as well. Furthermore, the inherent independence assumed for each centroid measurement allows errors to be measured not only in total, but also by position or by axis. In other words, the degradation of spatial linearity can be characterized by region within the FOV or by transaxial, sagittal, and coronal planes simply by observing grids points with the largest deviations from the synthetic coordinates.

7   Conclusion

Performance metrics for the limited angle tomography system were performed and revealed a number of limitations. Spatial resolution was generally found to be acceptable while deadtime was somewhat restrictive for practical applications. Other metrics including contrast-to-noise ratio (CNR) and resolution recovery were found to perform moderately well. Imaging artifacts, in particular, were prohibitively large for the sagittal and transaxial slices. The simple geometrical configuration of this system allows for a measure of parallax and intercrystal scatter that may not be easy to perform on a traditional cylindrical system. As a result, future work on performance should focus on this unique aspect of the limited angle tomography system.

In an attempt to compensate for this, Factor Analysis (FA) was applied to a number of imaging studies. A dual-isotope experiment proved unsuccessful, despite the fact that only two factors were required, and the fact that a limited number of voxels were selected for analysis. The results from the semi-quantitative study indicate that FA may be successful for sufficient counts. The mouse study indicated moderate success, but significant effort must, in general, be put forth in smoothing and reducing the data prior to applying FA. In short, FA does not appear to be a straightforward and simple method to counter the image artifacts from a limited tomography system. Future work may attempt to improve results by altering the choice of initial values, within the iterative simplex method used for the FA, in order to promote convergence.

These results prompted a more quantitative investigation into FA. The close relation between FA and Principal Component Analysis (PCA) allows for such an analysis without any loss of accuracy. Studies which attempted to measure the success of FA, by performing PCA, were conducted using a uniform cylindrical phantom at various activity levels and scan

durations. These results were used to provide an upper bound for FA. However, additional information was mathematically deduced which allowed the results to be interpreted in terms of uniformity, signal-to-noise ratio (SNR), and noise distribution within the FOV. This method was applied to a conventional cylindrical microPET system as well as the limited tomography microPET system within their nominal activity ranges. As expected, uniformity and SNR increase with activity and scan duration. For the cylindrical microPET system, uniformity and SNR were found to increase quite rapidly with increased activity or scan duration initially, but then leveled off at higher ranges. The uniformity and SNR increases within the limited angle tomography system prove to be much less robust. The change in SNR between data sets, within a given system, was found to be in close agreement with that predicated by Poisson statistics. For the cylindrical system, the level of agreement appeared to grow as the activity or scan duration increased; the limited angle tomography system also agreed with the Poisson model, but the level of agreement seemed to be nearly consistent regardless of the activity level or scan duration. Optimal levels of activity and scan duration were calculated, whenever possible, using a simple criterion of "diminishing returns". Noise distribution within both systems agreed very well with the simple zero-mean model. For all results, only the first principal component was computed for the correlation-based and covariance-based data sets. No assumptions or approximations were made in the calculation of this component. However, due to the sizes of these data sets, an alternative method of computing these components was developed. Future work may apply this method to other systems, modalities, and may even recursively compute the additional components beyond the first.

Coregistration between imaging systems was also performed via a rigid body transformation method. The positioning precision study indicates that, at least mechanically, the

microCT and microPET bed adapters are quite stable and highly suitable for use with this coregistration method. Calibration data for calculation of this transformation was performed using the same spherical phantom with random placement within each field-of-view (FOV). Calibration data taken from the microCT, in particular, proved somewhat troublesome due to the ambiguity in calculating the centroid of the spherical phantom. Coregistration of the calibration images themselves proved generally successful. As a result, coregistration of subsequent mice images was successful as well. However, the small visual discrepancies seen in the coregistered calibration data also appeared in the coregistered mice images, indicating that the method is consistent from scan to scan. This is expected given the nature of the rigid body transformation. Future work will attempt to improve the accuracy of the calibration data with the construction of a calibration phantom, as well as oversampling images prior to coregistration in order to reduce the roundoff errors stemming from converting physical coordinates to voxel indices. Future work may also involve coregistration between other small animal imaging modalities using nearly identical methods.

The assumption of a rigid body used for coregistering purposes may also be used to determine the spatial linearity of a given system. For these purposes, rigid body coregistration was performed between a synthetic regularly-spaced grid and its measured image. Due to the insensitivity of the coregistration to the chosen coordinate systems used for each set of measurements, the final results are only dependent on the imaging parameters. The metric provided by the coregistration is the total mean squared error between each grid point and its imaged point. Two-dimensional grids were printed with radioactive ink. High-count scans were acquired on the limited angle tomography system. Results indicate a fairly high degree of spatial linearity within the central coronal slice. Future work should include similar measurements

157

within other coronal slices near the edges of the FOV as that is where significant spatial distortion is likely to occur. As an alternative to the 2D printed grid, a 3D grid-like phantom may prove useful and allow for a more comprehensive study. This method may also be applied to other modalities assuming that a suitable regularly-spaced grid can be manufactured for that particular modality.

In conclusion, a number of venues have been explored in order to characterize the system performance of the limited angle tomography PET system. The severe drawbacks of this system may prevent its full applicability. In the process of investigation, relatively straightforward methods to measure several measures of performance have been developed. Many of these methods apply not only to most PET systems, but also to modalities other than PET. In fact, some of these methods are nearly identical in terms of mathematical foundation. While limited angle tomography may not be suitable for PET applications, it is hoped that these other methods will be very useful for many other imaging applications.

References

[1] H. Zhang, N. T. Vu, Q. Bao, R. W. Silverman, B. N. Berry-Pusey, A. Douraghy, D. A. Williams, F. R. Rannou, D. B. Stout and C. A. F, "Performance Characteristics of BGO Detectors for a Low Cost Preclinical PET Scanner," *IEEE Transactions on Nuclear Science,* vol. 57, no. 3, pp. 1038-1044, June 2010.

[2] H. Zhang, N. Vu, R. Silverman and R. Taschereau, "Performance Evaluation of PETbox: A Low Cost Bench Top Preclinical PET Scanner," *Molecular Imaging and Biology,* September 2010.

[3] "National Electrical Manufacturers Association Standards Publications NU 4: Performance Measurements of Small Animal Positron Emission Tomographs," National Electrical Manufacturers Association, Rosslyn, VA, 2008.

[4] C. C. Constantinescu and J. Mukherjee, "Performance Evaluation of an Inveon PET preclinical scanner," *Physics of Medicine and Biology,* vol. 54, pp. 2886-2898, 2009.

[5] Q. Bao, D. Newport, M. Chen, D. B. Stout and A. F. Chatziioannou, "Performance Evaluation of the Inveon Dedicated PET Preclinical Tomograph Based on the NEMA NU-4 Standards," *Journal of Nuclear Medicine,* vol. 50, pp. 401-408, 2009.

[6] S. R. Cherry, J. A. Sorenson and M. E. Phelps, Physics in Nuclear Medicine, 3rd ed., Elsevier Science, 2003.

[7] G. F. Knoll, Radiation Detection and Measurement, 3rd ed., Danvers, MA: John Wiley and Sons, Inc., 2000.

[8] D. Barber, "The use of principal component in the quantitative analysis of gamma camera dynamic studies," *Physics in Medicine and Biology,* vol. 25, no. 2, pp. 283-292, 1980.

[9] R. Di Paola, J. P. Bazin, F. Aubry, A. Aurengo, F. Cavailloles, J. Y. Herry and E. Kahn, "Handling of Dynamic Sequences in Nuclear Medicine," *IEEE Transactions on Nuclear Science,* vol. 29, no. 4, pp. 1310-1321, August 1982.

[10] C. Schiepers, C. K. Hoh, J. Nuyts, M. Seltzer, C. Wu, S.-C. Huang and M. Dahlbom, "1-11C-Acetate Kinetics of Prostate Cancer," *Journal of Nuclear Medicine,* vol. 49, pp. 206-215, February 2008.

[11] C. Schiepers, C. K. Hoh, J. Nuyts, H.-M. Wu, M. E. Phelps and M. Dahlbom, "Factor Analysis in Prostate Cancer: Delineation of Organ Structures and Automatic Generation of In- and Output Functions," *IEEE Transactions on Nuclear Science,* pp. 2338-2344, 2002.

[12] J. A. Nelder and J. Mead, "A simplex method for function minimization," *The Computer Journal,* vol. 7, no. 4, 1965.

[13] P. Heagerty, T. Lumley, G. van Belle and L. Fisher, Biostatistics: A Methodology for the Health Sciences, Wiley-Interscience, 2004.

[14] I. Griva, S. G. N. Nash and A. Sofer, Linear and Nonlinear Optimization, 2nd ed., Philadelphia, PA: Society for Industrial and Applied Mathematics, 2009.

[15] A. N. Gorban, B. Kegl, D. C. Wunsch and A. Zinovyev, Eds., Principal Manifolds for Data Visualization and Dimension Reduction, Springer, 2008.

[16] P. Sanguansat, Ed., Principal Component Analysis - Engineering Applications, InTech, 2012.

[17] O. Bretscher, Linear Algebra with Applications, Upper Sadle River, NJ: Prentice-Hall Inc., 1997.

[18] P. R. Bevington and D. K. Robinson, Data Reduction and Error Analysis for the Physical Sciences Second Edition, 2nd ed., McGraw-Hill Companies, 1992.

[19] C. Ji, F. van der Have, H. G. van Andel, R. Ramakers and F. Beekman, "Accurate Coregistration between Ulta-High Resolution Micro-SPECT and Circular Cone-Beam Micro-CT Scanners," *International Journal of Biomedical Imaging,* vol. 2010, July 2010.

[20] P. M. Bloomfield, T. J. Spinks, J. Reed, L. Schnorr, A. M. Westrip, L. Livieratos, R. Fulton and T. Jones, "The design and implementation of a motion correction scheme for neurological PET," *Physics in Medicine and Biology,* vol. 48, pp. 959-978, 2003.

[21] J. Qi and R. H. Huesman, "List mode reconstruction for PET with motion compensation: a simulation study," in *IEEE International Symposium on Biomedical Imaging Proceedings*, 2002.

[22] A. Rahmina, Ju-Chie, K. Dinelleb, M. S. P. W. Shilova, O. G. Rousseta, B. M. Tsuia, D. F. Wonga and V. Sossib, "System matrix modelling of externally tracked motion," *Nuclear*

Medicine Communiciations, vol. 29, pp. 574-581, November 2007.

[23] B. Horn, "Closed-form Solution of Absolute Orientation Using Orthornomal Matrices," *Journal of the Optical Society of America,* vol. 5, no. 7, pp. 1127-1135, March 1988.

[24] B. Horn, "Closed-form Solution of Absolute Orientation Using Unit Quaternions," *Journal of the Optical Society of America,* vol. 4, pp. 629-641, 1987.

[25] J. Kuipers, Quaternions and Rotation Sequences: A Primer with Applications to Orbits, Aerospace, and Virtual Reality, 5th ed., Princeton, NJ: Princeton University Press, 1999.

[26] B. H. Brown, R. H. Smallwood, D. C. Barber, P. V. Lawford and D. R. Hose, Medical Physics and Biomedical Engineering, 1999: Institute of Physics Publishing Ltd..

[27] M. J. Berger, J. S. Coursey, M. A. Zucker and J. Chang, "ESTAR, PSTAR, ASTAR: Computer Programs for Stopping-Power and Range Tables for Electrons, Protons, and Hellium Ions (version 1.2.3)," National Institute of Standards and Technology, Gaithersburg, MD, 2005. [Online].