"Positron emission tomography (PET) is the most important advance in biomedical science since the invention of microscope[1]"

Henry N. Wagner, Jr. M.D.
Johns Hopkins Bloomberg School of Public Health

University of Alberta

Target Volume Delineation in Dynamic Positron Emission Tomography Based on Time Activity Curve Differences

by

Artur Teymurazyan

A thesis submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Medical Physics

Department of Physics

Spring 2013
Edmonton, Alberta

To my parents Gayane and Rafael,
whose strength and love continue to inspire me
and
To my brother Aram, who is always there for me

To my wife Lyubov and children Alexander and Gregory,
I can not imagine a world without you

To Dr. Ani Aprahamian, a friend and mentor

To all my family and friends, with deepest love and respect

Abstract

Tumor volume delineation plays a critical role in radiation treatment planning and simulation, since inaccurately defined treatment volumes may lead to the overdosing of normal surrounding structures and potentially missing the cancerous tissue. However, the imaging modality almost exclusively used to determine tumor volumes, X-ray Computed Tomography (CT), does not readily exhibit a distinction between cancerous and normal tissue. It has been shown that CT data augmented with PET can improve radiation treatment plans by providing functional information not available otherwise.

Presently, static PET scans account for the majority of procedures performed in clinical practice. In the radiation therapy (RT) setting, these scans are visually inspected by a radiation oncologist for the purpose of tumor volume delineation. This approach, however, often results in significant interobserver variability when comparing contours drawn by different experts on the same PET/CT data sets. For this reason, a search for more objective contouring approaches is underway.

The major drawback of conventional tumor delineation in static PET images is the fact that two neighboring voxels of the same intensity can exhibit markedly different overall dynamics. Therefore, equal intensity voxels in a static analysis of a PET image may be falsely classified as belonging to the same tissue. Dynamic PET allows the evaluation of image data in the temporal domain, which often describes specific biochemical properties of the imaged

tissues. Analysis of dynamic PET data can be used to improve classification of the imaged volume into cancerous and normal tissue.

In this thesis we present a novel tumor volume delineation approach (Single Seed Region Growing algorithm in 4D (dynamic) PET or SSRG/4D-PET) in dynamic PET based on TAC (Time Activity Curve) differences. A partially-supervised approach is pursued in order to allow an expert reader to utilize the information available from other imaging modalities routinely used in conjunction with PET. In our scheme, this includes the definition of a tumor encompassing mask and selection of a seed site within the suspected tumor, while further delineation is performed automatically by the algorithm.

The development of this method is examined and improved classification of the imaged volume into cancerous and normal tissue compared to methods currently used in the clinic is demonstrated.

Acknowledgement

First and foremost, I would like to express my deepest appreciation to my thesis supervisor Dr. Don Robinson for introducing me to the PET research, for his encouragement and support of my work.

I express my gratitude to my co-supervisor Dr Terence Riauka for his input and suggestions, for taking the time to explain the practical aspects of PET imaging.

Special thanks go to Dr. Hans-Sonke Jans for his interest in my research and for being so accessible and helpful. He has been such a good person to discuss ideas with.

I would also like to extend my gratitude to Dr. Ron Sloboda for his kind attention and insightful discussions, and for the guidance he provided during this research project.

I am grateful to Dr. John Amanie for his clinical expertise and for providing us with clinical data and to the staff of Edmonton PET Centre and Nuclear Medicine Department of Cross Cancer Institute for preparation of the isotopes and the help with acquiring the PET data.

I would also like to thank all the people at Medical Physics Department of Cross Cancer Institute for always being helpful answering questions and for sharing their knowledge.

My sincere gratitude and love go to my beautiful, wonderful wife and best friend, Dr. Lyubov Titova. This dissertation would not have been completed without her endless love, encouragement and infinite patience through all these years.

Thank you all

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ACD	Annihilation Coincidence Detection
ACF	Attenuation Correction Factor
APD	Avalanche Photo Diode
AQ	Absolute Quantification
AW-OSEM	Attenuation-Weighted Ordered Subset Expectation Maximization
BLS-GSM	Bayes Least Squares - Gaussian Scale Mixture
BTV	Biological Target Volume
CDF	Cumulative Distribution Function
COV	Contour Overlap Fraction
CT	X-ray Computed Tomography
CT-3DAC	CT-based Three-Dimensional Attenuation Correction
CTV	Clinical Target Volume
CTW	Coincidence Timing Window
DWT	Discrete Wavelet Transform
FBP	Filtered Back-Projection
FDG	Fluoro-Deoxy-Glucose (^{18}F) or 2-deoxy-2-(^{18}F)fluoro-D-glucose
FOV	Field of View
FWHM	Full Width at Half-Maximum
GTV	Gross Tumor Volume
LLD	Lower Level Discriminator
LOR	Line of Response
MCP	Micro-Channel Plate (dynode)
MC-PMT	Multi-Chanel Photo-Multiplier Tube
ML-EM	Maximum-Likelihood Expectation Maximization
MR	Magnetic Resonance Imaging
MRI	Magnetic Resonance Imaging
MSE	Mean Squared Error
NSCLC	Non-Small Cell Lung Cancer
NSD	Normalized Standard Deviation
OSEM	Ordered Subset Expectation Maximization

PCA	Principal Component Analysis
PET	Positron Emission Tomography
PHA	Pulse-Height Analyzer
PMCC	Pearson's Product - Moment Correlation Coefficient
PMT	Photo-Multiplier Tube
PSNR	Peak Signal-to-Noise Ratio
PS-PMT	Position Sensitive Photo-Multiplier Tube
PTV	Planning Target Volume
QE	Quantum Efficiency
RAMLA	Row-Action Maximum-Likelihood Algorithm
RMSE	Root Mean Square Error
ROI	Region of Interest
RT	Radiation Therapy
SHINE	Statistical and Heuristic Image Noise Extraction
SNR	Signal to Noise Ratio
SPECT	Single-Photon Emission Computed Tomography
SSRG	Single Seed Region Growing
SSS	Single-Scatter Simulation
STD	Standard Deviation
SUV	Standardized Uptake Value
SUV (BW)	SUV normalized to the Body Weight
TAC	Time Activity Curve
TOF	Time-of-Flight
ULD	Upper Level Discriminator
VOI	Volume of Interest

1. Introduction

Positron Emission Tomography (PET) plays an ever increasing role in radiation therapy. In oncology, X-ray Computed Tomography (CT) (or Magnetic Resonance Imaging (MR or MRI)) data augmented with PET can improve radiation treatment plans by providing physiological function information not available otherwise[2-4]. For example, a potential increase of 14.3 % in tumor control probability has been demonstrated for treatment plans modified using dose-escalation factors calculated from dynamic [^{18}F]-fluoromisonidazole (FMISO) positron emission tomography data in hypoxia dose painting studies[5, 6]. Furthermore, PET/CT data has been proven to have a greater sensitivity and specificity in the staging of lung cancer[7] than either CT or MR alone. In the following sections a brief introductory review is presented for the reader to comprehend the goals, main contributions and findings in this work.

1.1 Motivation

In its early days, PET was mainly used as a research tool, however, in the last decade the role of PET in the clinic has become increasingly important. The largest area of clinical use of PET is in oncology. PET provides a unique tool for the visualization of biologic processes which can reveal valuable information pertinent to patient diagnosis, staging, progression and treatment outcome. The quantitative interpretation of PET images is, unfortunately, not always straightforward. Two confounding factors in the clinical application of PET are image noise and the interpretation of static image intensity values.

Radiation therapy techniques such as three-dimensional conformal radiation therapy (3D-CRT) and intensity modulated radiotherapy (IMRT) allow delivering large doses of radiation to small volumes. The ultimate goal is to improve local control by escalating dose to sites of disease. To achieve this goal, two potentially conflicting objectives must be met: (1) a target volume must be defined encompassing the entire tumor volume plus any geometric motions and uncertainties, (2) the volume of normal tissue within this defined target must be minimized. Accurate delineation of the full extent of malignant growth is crucial

in radiation oncology. An inaccurately defined target volume may potentially lead to underdosing the tumor or overdosing normal structures surrounding the tumor and thus jeopardize patient outcome.

Tumor localization has historically relied on the use of anatomic imaging modalities such as CT and MRI. Currently, CT based volume definition remains a gold standard for most radiotherapy treatment planning. CT acquisition and reconstruction parameters are standardized allowing the generation of consistently accurate, high resolution anatomical patient images. The intensity values in these images are reported in Hounsfield units (HU), which correlate with tissue density. The electron densities (related to tissue densities) of imaged tissues are required by modern treatment planning systems for correcting dose calculation in order to account for heterogeneities in patients. Unfortunately the change in Hounsfield units between normal and cancerous tissue is usually not very great, making accurate differentiation between normal and malignant tissue challenging.

The early reports on use of PET in radiation therapy were published in the late 1990s[8]. Recent technological advances in radiation therapy and multimodal imaging (PET/CT & PET/MR) have further stimulated the use of PET-guided radiation therapy treatment planning. In fact, there is an ongoing discussion regarding whether PET/CT has a potential to replace CT as a standard for radiotherapy simulation and planning. The main argument against is that even though PET has a potential to enhance the consistency of tumor volume definition, it lacks accuracy[9]. While CT (or MR) allows defining a "macroscopic" target volume with anatomic data, functional imaging modalities such as PET allow the extraction of another volume, a ''biological" target volume. The notion of a ''biological" target volume (BTV), in addition to the widely accepted concepts of gross tumor volume (GTV) and clinical target volume (CTV), led to the integration of tumor biology into the treatment planning process[8]. The rationale for the use of PET/CT in radiation therapy treatment planning is that in some cases CT fails to detect lesions that are readily visible on PET images. In addition, there are reports of smaller lesion volumes delineated on fused PET/CT when compared to volumes based on CT alone. Also, articles featuring the

disparity between volume delineation based on anatomical information only (CT, MR) and that based on PET are frequently published in the literature[8, 10-22]. These discrepancies are due in part to a higher sensitivity and specificity of PET for the identification of tumor tissue in comparison to CT and MRI alone. Another limitation of using anatomical imaging modalities in RT treatment planning is a difficulty in visualization and accurate delineation of tumors in postsurgical distorted anatomy and the presence of residual scar tissue after chemotherapy[23].

FDG (^{18}F-fluoro-deoxy-glucose) is the radiopharmaceutical used in PET imaging that has the widest application in clinical oncology. FDG is relatively easy to synthesize with a high radiochemical yield[24]. A typical activity of FDG used in an oncologic scan is 200-400 MBq for an adult human. This tracer is glucose analog and follows a similar metabolic pathway. Because an oxygen atom is replaced by ^{18}F, FDG is not metabolized to CO_2 and water and remains trapped within tissue until it decays. The result of this is intense radiolabeling of tissues with high glucose uptake, such as the brain, the liver, and most cancers[25]. Oncologic scans using FDG make up over 90% of all PET scans in current practice[26]. Various authors have published reports on the feasibility of FDG-PET/CT based tumor volume delineation in head and neck tumors, esophageal cancer, NSCLC, breast cancer, cervical cancer, Hodgkins lymphoma, and rectal cancer[8]. Novel probes beyond FDG for delineation of other tumors and application in RT treatment planning are constantly emerging[8].

The most straightforward integration of the functional information provided by PET imaging into the RT treatment planning process consists in (i) delineating tumor volumes on a PET study, and (ii) transferring these volumes on CT images after registration. The image registration step has not only been well studied[27-31], but also has been significantly simplified by multimodal imaging (PET/CT & PET/MR). On the other hand, the accurate delineation of PET defined targets remains a problem[16, 27, 32] in spite of the vast literature accumulated on this topic over the years[33]. For the many methods currently used in clinical practice, delineation results can be quite disappointing especially in the case of small heterogeneous or non-spherical tumors[33].

In this thesis we present a novel single seed region growing algorithm for tumor segmentation in dynamic positron emission tomography (henceforth SSRG/4D-PET). We compare SSRG/4D-PET to current clinical tumor segmentation methods and demonstrate improved classification of an imaged volume into cancerous and normal tissue may be achieved in the presence of both homogeneous and heterogeneous uptake.

1.2 Thesis Objectives

The main goal of research described in this thesis is the development of a novel partially-supervised tumor delineation algorithm incorporating both temporal and spatial information available in dynamic PET studies for use in radiation therapy treatment planning. The vision is to improve delineation accuracy of the geometric extent of PET defined target volumes by analyzing and exploiting the differences in the time activity curves (TAC) for tumor and surrounding normal tissue.

To reach the aims of this work, first, statistical properties of noise in PET images reconstructed either analytically or iteratively are studied. This has important implications for selection of an appropriate noise suppression technique prior to delineation as well as for simulation of noise in synthetic PET data used in delineation algorithm performance evaluation. Following this, three different denoising approaches (PCA, BLS-GSM, and SHINE) are compared and the most appropriate noise suppression technique (PCA) is integrated into the framework of the proposed algorithm (SSRG/4D-PET). The algorithm performance is evaluated using semi-empirical digital phantoms (homogeneous and heterogeneous), a clinically-relevant physical phantom, and clinical patient images.

1.3 Thesis Contributions

The following summarizes the contributions that have been made as a result of this thesis:

- An examination of the statistical properties of noise in reconstructed PET images after all the necessary corrections has been applied. Statistical properties of the acquired data are evaluated and compared to five noise models (Poisson, Normal, Negative Binomial, Log-Normal and Gamma). This analysis shows that the noise in RAMLA reconstructed PET images is very well characterized by Gamma distribution followed closely by Normal distribution, while FBP produces comparable conformity with both Normal and Gamma statistics. These results have been published in Journal of Digital Imaging (*Teymurazyan et al*. 2012)[34].

- Three different approaches (PCA, BLS-GSM, and SHINE) to noise suppression in reconstructed PET images are evaluated. Their performance is compared to some recently published state-of-the-art methods as well as a number of standard well-known and widely-used denoising algorithms.

- A comparison between denoising PET images with different techniques reveals PCA to be an appropriate method for post-processing of dynamic PET data aimed at noise reduction prior to tumor volume delineation. Within the framework of the proposed algorithm, retaining only the first principal component (PC1) allows obtaining a data set with a much lower level of noise while enhancing contrast, thereby optimizing the signal-to-noise ratio (SNR).

- A novel single seed region growing algorithm for dynamic positron emission tomography (henceforth SSRG/4D-PET) is presented. We demonstrate that improved classification of the imaged volume into cancerous and normal tissue can be achieved compared to other methods currently used in the clinic. These results have been published in IEEE Transactions on Nuclear Science (*Teymurazyan et al*. 2012)[35].

1.4 Thesis Outline

In *Chapter 2*, background information on Positron Emission Tomography (PET) is presented. This begins with an overview of PET technology and scanner detection mechanisms, followed by a brief discussion of image acquisition, image reconstruction, and the necessary corrections which are critical for conversion of raw measurements into a time series of three-dimensional images. The chapter concludes with a section on absolute quantification, which is used to assess patient images in order to distinguish benign tissue from malignant and (or) to assess response to therapy.

Chapter 3 describes the properties of noise in reconstructed PET images, which is one of the confounding factors in PET imaging. This chapter examines the statistical properties of noise in PET images acquired with a GEMINI PET/CT scanner and reconstructed with Filtered-Backprojection (FBP) and Row-Action Maximum Likelihood Algorithm (RAMLA), after all clinical correction and image reconstruction procedures have been applied. This has important implications for an accurate evaluation and processing of quantitative information provided by PET imaging. Effective image noise reduction is greatly dependent on an accurate knowledge of the parameters which characterize this noise. A version of this chapter has been published in Journal of Digital Imaging (*Teymurazyan et al*. 2012)[34].

Chapter 4 focuses on image denoising. Since noise levels observed in PET images complicate their geometric interpretation, it is desirable to implement an effective post-processing technique aimed at noise reduction prior to tumor volume delineation. This chapter briefly discusses and compares three different approaches (PCA, BLS-GSM, and SHINE) to noise suppression in reconstructed PET images.

Chapter 5 summarizes the segmentation techniques used in medical imaging. A review of the image segmentation literature is presented, including the methods currently used for segmentation of PET data. We first present current practices for tumor volume delineation and discuss their limitations. We then demonstrate how some of these shortcomings can be overcome with dynamic (4D) PET imaging.

Chapter 6 presents a single seed region growing algorithm for tumor volume delineation in dynamic PET imaging (SSRG/4D-PET). After fully describing the algorithm followed by a description of methods and various figures of merit used, an in depth evaluation of the proposed methodology is presented. A version of this chapter has been published in IEEE Transactions on Nuclear Science (*Teymurazyan et al*. 2012)[35].

Finally, *Chapter 7* summarizes the thesis and presents the conclusions drawn as a result of this study, after which future work directions are discussed for further improvement of tumor volume delineation in PET.

2. Positron Emission Tomography in Medical Imaging

Positron Emission Tomography (PET) is a non invasive, medium resolution, diagnostic imaging modality. This technique is based on the decay of a positron (β^+) emitting radionuclide which is chemically bound to a biologically active molecule (tracer). Of particular interest are the radioactive forms of biologically important elements such as oxygen, nitrogen, carbon and fluorine, allowing in-vivo imaging of physiological function with PET. The advantages over other imaging modalities used in Nuclear Medicine are increased sensitivity and the possibility of accurate attenuation correction. PET is capable of detecting very small (picomolar) quantities of radioactive agents administered to patients[36]. The radiation detected by a PET scanner is used to map the three dimensional (3D) distribution of radiopharmaceuticals in the body. Time sequences can be acquired (4D PET) to determine the time evolution of activity, with the aim of extracting physiological constants such as metabolic activity or transport rates[36]. PET imaging is relatively expensive as it requires not only a PET scanner but typically also a cyclotron for tracer production due to the short half-life of the radioactive materials involved. On the other hand PET has unique functional capabilities and a wide ranging research and clinical potential. It is used worldwide as a diagnostic tool in oncology, cardiology, neurology and pharmacology. With the advent of PET/CT, its role in oncology in general and in radiation therapy in particular is increasing.

PET is unique, because it is able to visualize biochemical processes altered by disease and can detect these alterations before there is an actual change in patient anatomy. This separates it from traditional diagnostic techniques, such as x-rays, CT scans, or MRI, which produce images of the body's anatomy or structure. Even in diseases such as Alzheimer's, where there is no anatomical abnormality, PET can reveal changes in biochemistry. In clinical oncology it is useful in managing patients with certain conditions affecting the brain and the heart as well as with certain types of cancer.

2.1 Sources of Positrons

A positron (β^+ particle) is the antimatter conjugate of an electron emitted in beta decay. Its existence was first postulated by Dirac around 1928 and was later discovered by Anderson in 1832. Positrons are produced by either pair production or in the course of nuclear transmutation called beta-plus decay (weak nuclear decay). The radioisotopes used in PET are proton-rich and decay via positron emission, in which a proton in the nucleus is converted into a neutron (^1_0n), a positron ($^0_1\beta^+$) and a neutrino (ν_e).

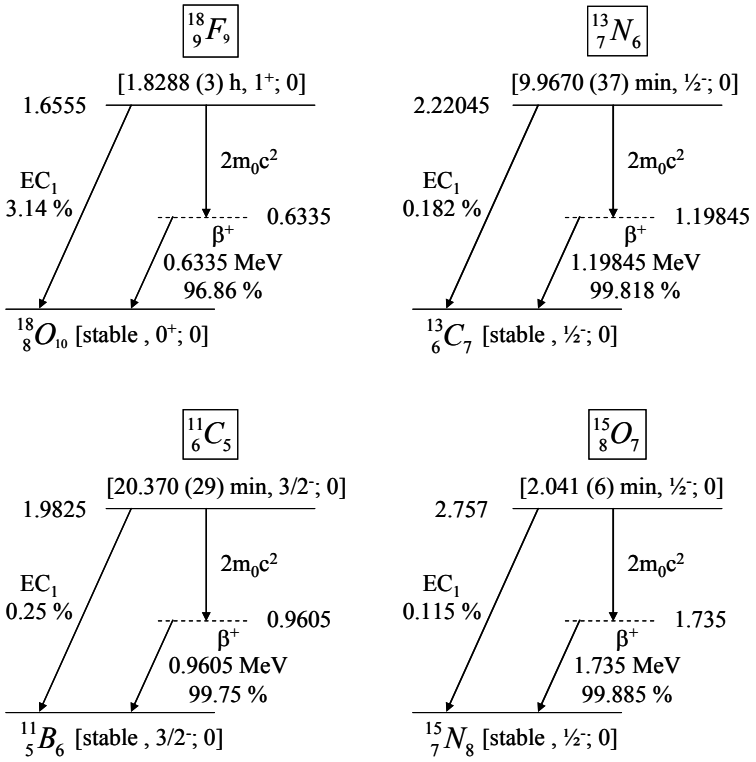

Figure 2.1: Decay scheme of common β^+ emitters

In general, the positron decay from the atom can be expressed as[37]:

$$^A_ZX \rightarrow {}^A_{Z-1}X + {}^0_1\beta^+ + \nu_e + Q(+e^-) \qquad \textbf{2.1}$$

where Q is energy. The daughter isotope has an atomic number of one less than the parent, since positive charge is carried away with the positron. To balance the

charge of the daughter nucleus, an orbital electron must be ejected from the atom, often via the process known as internal conversion. The nucleus supplies the energy required by the orbital electron to overcome its binding energy and leave the atom with residual kinetic energy. In order for an atom to undergo positron decay its atomic mass has to exceed the mass of the daughter by at least two electron masses ($2m_0c^2$), because both a positron and an electron are emitted in the process. **Figure 2.1** shows beta-plus decay schemes of ^{18}F, ^{11}C, ^{15}O, and ^{13}N. The vertical distance between the horizontal solid lines in **Figure 2.1** represents the total energy released (the transition energy) in the decay process and is also known as the *Q value* (**2.1**).

2.2 PET Imaging

PET imaging relies on the tracer principle: radiopharmaceuticals are distributed, metabolized, and excreted according to their chemical structure. The radiotracer is introduced into a patient usually via injection. Depending on radionuclide half-life and on the specific type of PET examination, a waiting period (normally between 30 min to 2 hours) is required between tracer injection and scanning to allow the metabolically active molecule to become concentrated in tissues of interest. The short-lived proton-rich isotope decays, emitting a positron. The positron produced travels for a short distance (~1 mm) through surrounding tissue, losing kinetic energy principally by Coulomb interactions with electrons, as it collides with molecules.

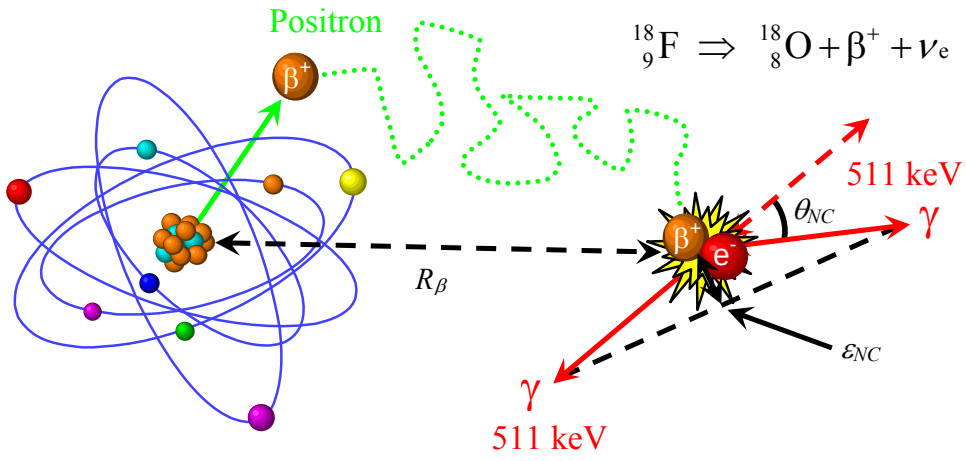

Figure 2.2: Positron emission and annihilation process

Since the rest mass of a positron is equal to that of an electron, positrons may undergo large deviations in direction with each Coulomb interaction, thus following a very complex path through tissue (**Figure 2.2**). This path can be much longer than the Euclidean distance (positron range, R_β) between point of emission and point of annihilation of the positron. Because the centre of mass of a positron and an atomic electron is not always at rest at annihilation, conservation of energy and momentum dictates that the annihilation photons produced are not always emitted at 180° from each other. This deviation from colinearity, which is shown on **Figure 2.2** as an angle of non-colinearity (θ_{NC}), results in an annihilation localization error (ε_{NC}). This error arises because the detection process in PET assumes that annihilation photons travel in exactly opposite directions along a straight line called the line of response (LOR), shown as a black dashed line between the gamma rays in **Figure 2.2**. The effect of *NC* on PET image resolution will be discussed later (see section **2.5**).

2.3 PET Radionuclides

One of the significant reasons for the importance of PET in medical research and in the clinic is the existence of positron-emitting isotopes which may be used to create a range of tracer compounds with properties similar to those of naturally

occurring substances in the body. This means that suitably labeled compounds can be introduced into the body without significantly altering its normal processes. An ideal PET tracer has to satisfy to two conditions: 1) be readily available or relatively simple to produce; 2) be suitable for the synthesis of biological compounds used in vivo for studies of biochemical processes. Physical characteristics of the current principal PET radionuclides are presented in **Table 2.1**.

Table 2.1: Common PET isotopes and their physical characteristics

Radionuclide	Source	β+ decay branching fraction (%)	Half-life (min)	Max (Mean) positron energies (keV)	Max (Mean) range in water (mm)
C-11	Cyclotron	99.75	20.37	961 (390)	4.1 (1.1)
N-13	Cyclotron	99.818	9.97	1198 (490)	5.4 (1.3)
O-15	Cyclotron	99.885	2.04	1735 (740)	7.3 (2.5)
F-18	Cyclotron	96.86	109.73	634 (250)	2.4 (0.6)
Ga-68	Generator	89.14026	67.71	1899 (836)	8.2 (2.9)
Rb-82	Generator	95.5	1.25	3356 (1532)	14.1 (5.9)

These radionuclides are incorporated either into compounds normally used by the body such as glucose (or glucose analogues), water or ammonia, or into molecules that bind to receptors or other sites of drug action. Since most radionuclides used in PET are relatively short lived (see **Table 2.1**), they are usually produced on site by means of dedicated medical cyclotrons or generators (for example ^{68}Ga and ^{82}Rb). ^{18}F is a noteworthy exception; with a half-life of almost 110 min it can be distributed to smaller facilities at a reasonable distance from its production site.

2.4 Coincidence Detection and Electronic Collimation

As mentioned, PET involves the detection of two photons traveling in opposite directions in coincidence. After injection of a tracer compound labeled with a positron emitting radionuclide, the patient is placed within the field of view (FOV) of a number of detectors capable of registering incident gamma rays. In current scanners, one or more rings of small scintillation detectors are typically

used. These detectors convert high energy annihilation photons into lower energy photons in the visible and ultraviolet range of wavelength. These lower energy photons interact with a photodetector which converts them into an electrical signal. The reshaped and amplified signal is accepted by energy discrimination circuitry if it falls within a predetermined energy window (~350-650 keV[38]). Finally, each detector generates a timed pulse when it registers an incident photon. When two pulses from opposing detectors arrive within a short time window[a], an event (prompt) is registered for those detectors. The width of the coincidence timing window is determined by the geometry of the scanner and is typically taken to be slightly longer than the time required for the photon to cross the ring. The two coincidence photons detected in a PET scanner may originate from anywhere within the scanner's FOV; thus it is possible for each photon to travel a different distance prior to interaction within the detectors. Theoretically this distance can be as large as the scanner's diameter. Using the diameter of a typical whole-body scanner (100 cm) and the speed of light ($c = 3\times10^8$ m/s), one can calculate the minimum requirement for the width of the coincidence timing window (~3-4 ns)[39]. This places a lower limit on the size of the coincidence timing window due to the difference in arrival times of two photons emitted at the edge of the scanner FOV. In realty, with the possible exception of very fast scintillators such as BaF_2 (scintillation decay time of 0.6 ns), the timing resolution of the detectors needs to also be accounted for when defining the coincidence timing window. For detectors with poor timing resolution, a large coincidence timing window is required in order to detect the majority of valid coincidence events[39]. Near-simultaneous detection of two annihilation photons in opposing detectors, also known as annihilation coincidence detection (ACD), actually defines the volume from which they are emitted. It allows PET to localize the origin of the annihilation event along a line (line of response or LOR) between

[a] This window duration is about 12 ns for a typical bismuth germinate (BGO) scanner (P. E. Valk, Positron emission tomography: basic science and clinical practice, edited by P. E. Valk (Springer, London ; New York, 2003), pp. 884).

two the detectors. ACD allows obtaining positional information without the need of mechanical[b] collimation. This mechanism is called electronic collimation. Electronic collimation results in an increased sensitivity (two to three orders of magnitude higher) and improved uniformity for PET compared to SPECT[c], where a mechanical collimator is used to restrict acceptance angles. The period of time during which events are collected is known as the scan time. For each pair of detectors, the total number of events registered equals the sum of the distribution of the activity or line integral along the LOR. The acquisition system reconstructs the acquired information into medical images. Position resolution within reconstructed images and the reconstruction process itself are some of the key features of PET imaging which are limited by several physical and technological factors, which shall be dealt with subsequently.

2.5 The Fundamental Limits of PET

In general, the resolution of a PET system is not shift-invariant and should be treated as a three dimensional parameter[40-42]. It is common to describe overall PET system resolution in terms of transaxial (in-plane) and axial components. The resolution of a PET system is also dependent on the mode (2D or 3D, see section **2.7**) of operation and is normally measured in terms of the FWHM of a small (≤ 1 mm in diameter) point source of activity[41]. There are four major independent factors affecting the point source spatial resolution of a PET system. Theoretical limits for the obtainable spatial resolution (R_{image}) of a PET system can be estimated if certain factors are assumed to be isotropic in the transaxial and axial directions[40]. From a physics point of view, these blurring factors are the positron range (R_{β}), annihilation non-colinearity (R_{NC}), and detector size ($R_{\text{det-}size}$). From the technological point of view a major contributing factor, for scanners with detector elements organized in blocks (see section **2.6.3**), is the block effect

[b] This is an absorptive collimation technique used in other nuclear imaging modalities.

[c] SPECT: Single-Photon Emission Computed Tomography. The technetium-99m (^{99m}Tc) emitting 140.5 keV gamma-rays is the most commonly used radionuclide in SPECT.

(R_{block}). The overall system resolution can be estimated using an expression determined empirically by Moses *et. al.*[42, 43]:

$$R_{image} = 1.25 \times \sqrt{R_{\beta}^2 + R_{NC}^2 + R_{det\text{-}size}^2 + R_{block}^2} \qquad \textbf{2.2}$$

where the factor 1.25 accounts for the effect of the reconstruction process.

The contribution from positron range to overall system resolution stems from the uncertainty in the exact location of positron emission. Positrons are emitted with a continuous spectrum of kinetic energy ranging from zero to a maximum energy (**Table 2.1**). Based on their initial kinetic energy and the electron density of the biological medium, positrons travel along a random trajectory so that the traveled path is much longer than the distance (R_{β}) between their point of emission and annihilation (**Figure 2.2**). Thus positron range is a distribution of values that is normally characterized by a full width at half-maximum (FWHM) value. The R_{β} in **2.2** is the average distance traveled by the positron before annihilation. For most isotopes of interest, the FWHM range is 0.5 - 2.5 mm (**Table 2.1**).

The annihilation photons are not always emitted at 180° from each other as positronium. This deviation is shown in **Figure 2.2** as the angle of non-colinearity (θ_{NC}). The quoted value for the angle of non-colinearity is $\theta_{NC} = 0.25°$ [40, 42]. The non-colinearity effect can be described as:

$$R_{NC} = 0.5 \times D \times \tan\left(0.25^{\circ}\right) = 0.0022 \times D \qquad \textbf{2.3}$$

where D (in millimeters) is the distance between coincident detectors or the diameter of the detection system. For the Philips Allegro/Gemini scanner D is equal to 800 mm, therefore the contribution of non-colinearity to the effective spatial resolution of the scanner is 1.76 mm.

Due to the finite dimensions of radiation detectors, the LOR is really a parallelepiped joining any two detector elements as a volume of response. For a point source at the centre of the volume of response, and a detector width much smaller than the system diameter, the detector size response ($R_{det\text{-}size}$) is given by a FWHM equal to half the detector element width in either the transaxial or axial

directions[40, 42]. For the Philips Allegro/Gemini scanner, the transaxial crystal size of 4 mm gives a detector element contribution to the resolution uncertainty of 4/2 = 2 mm.

The block effect is the detector positioning accuracy resulting from the crystal identification algorithm[42]. For scanners based on detector elements organized in blocks, R_{block} is equal to 2 mm.

Thus, the theoretical limit for the obtainable spatial resolution of the Philips Allegro/Gemini PET system (for ^{18}F) is equal to $R_{image} = 1.25 \times (0.6^2 + 1.76^2 + 2^2 + 2^2)^{1/2} = 4.23$ mm.

2.6 Detector Configuration and Crystals used in PET

The main criteria describing a good scintillator for radiation detection in PET are the stopping power for 511 keV photons, light yield, scintillation signal decay time, and intrinsic energy resolution. Other desirable qualities are low cost of manufacturing, radiation hardness (to minimize radiation induced changes in the properties of the crystal), ruggedness (ease of fabrication, less likely to break while in use), non-hygroscopic (eliminates need for hermetic seal) and durability of operational parameters. The characteristics of inorganic scintillators are discussed in section **2.6.1**. Section **2.6.2** describes typical crystalline scintillators used for PET and an overview of detector configurations used in PET is presented in section **2.6.3**.

2.6.1 Characteristics of inorganic scintillators

Coincidence detection of the energetic annihilation photons characteristic to PET places specific requirements on the scintillator materials used. An ideal scintillation crystal should have the following properties: high density, high atomic number (large photoelectric cross-section), short scintillation decay time (good coincidence timing), large light output per keV of photon energy (good energy resolution), narrow emission spectrum around 400 nm (good coupling to a PMT photocathode), and transparency to emission photons (low self-absorption).

To maximize the sensitivity of a PET scanner, it is desirable to select a scintillation crystal with high stopping power for 511 keV photons in order to maximize the number of photons interacting and depositing energy in the crystal. The stopping power of the detector is characterized by its attenuation length [42], defined as the mean distance traveled by a photon before it interacts with the crystal:

$$\text{attenuation length} = 1/\mu \qquad \textbf{2.4}$$

where μ refers to the linear attenuation coefficient of the material (*units:* cm^{-1}). The linear attenuation coefficient (μ) is a measure of the probability that a photon will interact within a unit length of the material. Thus, a scintillation crystal with large linear attenuation coefficient (short attenuation length) at 511 keV is needed for a PET scanner with high efficiency. The linear attenuation coefficient is directly proportional to the total atomic cross-section for photon interactions, which is given as the sum of the cross-sections of all individual interaction mechanisms of the photons with matter. The main interaction mechanisms for annihilation photons in the 511 keV range are Compton scattering and photoelectric absorption. In Compton scattering within a scintillation crystal, only a portion of the energy of the annihilation photon is transferred to the scattered electron and the rest is taken by the scattered photon of lower energy, resulting in secondary Compton and photoelectric events. On other hand, in photoelectric absorption, the entire energy is converted to the release of a photoelectron which excites higher energy states of the crystal lattice, thereby producing a single scintillation center. Due to position blurring associated with Compton scattering within the crystal, photoelectric absorption is the preferred pathway. Photoelectric absorption is commonly parameterized as the photoelectric absorption probability (*PE*) or photofraction. The photofraction is defined as the probability that a photon will undergo photoelectric absorption instead of Compton scattering:

$$PE = 100\% \cdot \frac{\sigma_{ph}}{\sigma_{ph} + \sigma_C} \qquad \textbf{2.5}$$

where σ_{ph} and σ_c are photoelectric absorption and Compton scattering cross-sections. The Compton scattering cross-section is a linear function of electron density and thus proportional to the density of the crystal ρ[44], while the photoelectric absorption cross-section is related both to density (ρ) and the effective atomic number Z_{eff} describing the attenuation properties of a mixture of atoms in a molecule[42]. The effective atomic number is defined as:

$$Z_{eff} = \sqrt[\alpha]{\sum_{i=1}^{n} \frac{m_i Z_i}{\sum_{j=1}^{n} m_j Z_j} Z_i^{\alpha}} \qquad \textbf{2.6}$$

where m_i is the number of atoms of element i in the molecule. The photoelectric cross section is proportional to ρZ_{eff}^{α}, where α varies between 4 and 4.8 depending on the energy of the interacting photon. Thus, to maximize the fraction of photons undergoing photoelectric absorption, a scintillation crystal with high density and high atomic number is required.

The decay time of the excited state determines the timing characteristics of the scanner[42]. Longer decay times result in wider coincidence time windows. Wide coincidence time windows reduce the efficiency of the scanner by limiting performance at high count rates and increasing the number of random coincidences observed. High count rates limit the ability of the scanner to process each pulse individually leading to pulse pile-up and decreased sensitivity, if crystals with long decay times are used. Additionally, the ideal scintillator should have a short rise time associated with the luminescence process. Short rise and decay times are desirable for optimal timing resolution, since both of these characteristics determine the initial scintillation photon emission rate I_0[d].

High light output (light yield)[e] of the scintillator helps to achieve fast timing resolution, good spatial resolution with a high encoding ratio[f] and good energy

[d] Short rise time and decay times imply high initial scintillation photon emission rate.

[e] Number of emitted scintillation photons per MeV of energy absorbed.

resolution[42]. Eight to sixty light photons are emitted per *keV* of incident energy depending on the type of crystal used (**Table 2.2**). High light output results in good energy resolution, which is proportional to the square root of the number of scintillation photons. High light collection efficiency of the crystal and minimal optical self-absorption are imperative for the detection of a large fraction of emitted photons. The emission spectrum of the scintillator should overlap with the spectral sensitivity of the photodetector to ensure efficient conversion of captured light into a pulse. With the exception of the surface of the facing photodetector, the scintillator should be coated (or surrounded) with reflective material to prevent light from escaping the crystal. Furthermore, it is desirable for scintillation light output to be proportional to energy deposited in the crystal. Otherwise, the difference in scintillation light output due to full 511 keV energy absorption via a single photoelectric absorption event as compared to full energy absorption via multiple (lower energy) Compton interactions will lead to broadening of the full energy peak, thus degrading the energy resolution ($\Delta E/E$) of a PET detector.

Good energy resolution is required for effective rejection of events reaching the detector after undergoing Compton scattering inside the patient. This in turn affects the spatial resolution of the detector, since these events if not rejected are assigned to incorrect LORs which degrade spatial resolution. The energy resolution of a PET detector, besides the scintillation light output, also depends on the intrinsic energy resolution of the scintillator. Intrinsic energy resolution is affected by inhomogeneities arising in the crystal during the growth process as well as random variations in the production of light within it[38, 39].

[f] Encoding ratio is ratio of number of resolution elements, or crystals, to number of photodetectors.

Table 2.2: Characteristics of scintillator crystals used in PET (data from *Lecomte et al.*[42])

Properties	Typical inorganic scintillators								
	NaI	**BGO**	**GSO**	**LSO**	**LYSO**	**LGSO**	**LuAP**	**YAP**	**$LaBr_3$**
Peak emission wavelength (nm)	410	480	440	420	420	415	365	350	360
Index of refraction	1.85	2.15	1.85	1.82	1.81	1.8	1.94	1.95	1.9
Light yield (ph/keV)	41	9	8	30	30	16	12	17	60
$\tau_{s,}$ scintillation decay time (ns)	230	300	60	40	40	65	18	30	16
I_0 at 511 keV (ph/ns)	90	21	60	380	380	125	340	290	1,900
$\Delta E/E$ (%) at 662 keV	6	10	8	10	10	9	15	4.5	3
Density ρ (g/cm^3)	3.67	7.13	6.71	7.35	7.19	6.5	8.34	5.5	5.3
Effective Z (Z_{eff})	50	73	58	65	64	59	65	33	46
$1/\mu$ at 511 keV (mm)	25.9	11.2	15	12.3	12.6	14.3	11	21.3	22.3
PE (%)[a]	18	44	26	34	33	28	32	4.4	14
PE^2 (%)[b]	3.2	19	6.8	12	11	7.9	10	0.2	1.9
Hygroscopic	Yes	No	No	No	No	No	No	No	Yes
Magnetic	No	No	Yes	No	No	Yes	No	No	No

NaI thallium doped sodium iodide - *NaI*(*Tl*), *BGO* bismuth germanate ($Bi_4Ge_3O_{12}$), *GSO* cerium doped gadolinium orthosilicate - $Gd_2SiO_5(Ce)$, *LSO* cerium doped lutetium orthosilicate - $Lu_2SiO_5(Ce)$, *LYSO* cerium doped lutetium-yttrium oxyorthosilicate - $Lu_{1.9}Y_{0.1}SiO_5(Ce)$, *LGSO* cerium doped lutetium-gadolinium oxyorthosilicate - $Lu_{0.4}Gd_{1.6}SiO_5(Ce)$, *LuAP* cerium doped lutetium-aluminum perovskite ($LuAlO_3$:*Ce*), *YAP* cerium doped yttrium-aluminum perovskite - $YAlO_3(Ce)$, $LaBr_3$ cerium doped lanthanum bromide $LaBr_3(Ce)$

[a] Photoelectric absorption probability (or photofraction) at 511 keV

[b] Coincident photoelectric absorption probability

2.6.2 Typical crystalline scintillators for PET

The characteristics of typical crystalline scintillators are presented in **Table 2.2**. Thallium doped sodium iodide, NaI(Tl), is a reference scintillator widely used in SPECT for its excellent light yield, energy resolution and relatively high Z_{eff}. It has been used in PET scanners despite its lower stopping power for 511 keV photons and relatively slow decay time. Slow decay time and the necessity of using very thick crystals due to its low density limit attainable image resolution. Furthermore, this scintillator is neither rugged nor non-hygroscopic[g] making it vulnerable to moisture in the air as well as mechanical and thermal shocks, further reducing its value for PET applications.

The high density and high effective atomic number Z_{eff} of bismuth germanate (BGO) result in a very good stopping power for 511 keV gamma photons and a high photofraction *PE*. The photoelectric absorption cross section (σ_{ph}) of BGO at 511 keV is 1.6 times larger than for LSO and 5.8 times larger than that of NaI(Tl)[44]. Both BGO and NaI(Tl) are slow scintillators, i.e. both of these materials have relatively long decay constants. These longer decay constants result in increased dead time, thereby limiting the count rates at which the system may operate. Despite its low light yield and slow re-emission (visible light) of the energy absorbed in interactions with annihilation photons, BGO has been widely used in clinical PET scanners since the 1990's due to its high detection efficiency[45].

BGO is being phased out in most PET applications and replaced by materials (**Table 2.2**) doped with cerium in the trivalent charge state (Ce^{3+})[42]. Similar in density and Z_{eff} to BGO, but with a higher detection efficiency, initial scintillation photon emission rate I_0, photofraction *PE*, light yield, short decay time τ_s (**Table 2.2**) and a short rise time of 0.5 ns, make LSO/LYSO arguably the most suitable scintillator for PET and TOF-PET imaging currently available[42, 44]. One disadvantage of LSO/LYSO is the non-proportionality of scintillation light output to deposited energy[44] and the presence of the naturally radioactive lutetium

[g] NaI(Tl) requires careful handling and needs to be hermetically sealed with aluminum foil.

isotope. Naturally occurring lutetium consists of two isotopes: stable ^{175}Lu and radioactive ^{176}Lu (β^- decay) with 2.6 % abundance and a half-life of 4×10^{10} years. Single events in a PET detector due to the presence of this isotope have negligible impact at clinical count rates. These events may, however, have a noticeable effect in low count rate measurements such as research related imaging with dedicated small animal PET and some clinical quality control procedures[44, 45].

Gadolinium orthosilicate (GSO) is another interesting scintillator with properties useful as a PET detector. The advantages of this material are higher energy resolution than BGO and LSO, similar scintillation decay time to LSO and LYSO and more uniform light output than LSO. However, it has slightly lower stopping power and light output than LSO. Another disadvantage of GSO, and to a lesser degree LGSO, is the presence of large amounts of gadolinium. The high magnetic susceptibility of gadolinium and its widespread use as an MR contrast agent make GSO/LGSO unsuitable for use in combination with MR instruments[42].

Some other examples of scintillator materials potentially attractive for PET applications are lutetium aluminum perovskite (LuAP), yttrium-aluminum perovskite (YAP) and lanthanum bromide ($LaBr_3$). YAP has been used in small animal PET scanners[44, 45]. Despite the attractive properties of LuAP (see **Table 2.2**), its usefulness is limited due to restrictions on the thickness (less than a centimeter) of the crystal, arising from the strong self-absorption of scintillation light[44]. Lanthanum bromide is another good candidate for PET with excellent light output, the best initial scintillation photon emission rate I_0 of all known scintillators[42], highest reported energy resolution (<3% @ 662 keV), and timing properties which are suitable for TOF-PET measurements. Both LuAP and $LaBr_3$, like LSO, contain naturally occurring ^{176}Lu. The introduction of these and other new scintillating materials is limited due to cost and availability[44, 45].

2.6.3 Detector Configuration

To facilitate signal readout the scintillation crystals must be coupled to photodetectors. The size of these crystals determines the intrinsic resolution of the PET system[36]. Presently, a large number of small crystals, with cross-sectional dimensions of 4-8 mm (e.g. cross-sections of 4×4, 4×6, 6×6 and 4×8 mm^2) and depth of 20-30 mm, are used[45]. In general, the coupling of a scintillator to a photodetector is achieved in one of two ways: (1) large number of small crystal elements coupled onto photomultiplier tubes with larger areas; (2) one to one coupling, where a single crystal is paired with a single photo-detector.

The first of these (multiple crystals coupled to a single PMT) is used in the detector readout scheme commonly referred to as the block-detector principle, where light sharing and centroid calculation is employed to identify in which crystal of a matrix a gamma ray was detected[36, 39]. For example, in a single block, a matrix of 8×8 crystal elements can be encoded on an array of 2×2 PMTs[39, 45]. Round or square PMTs of 1-5 cm cross-section are commonly used in PET scanners[38]. Each block detector is about 3 cm thick and grooved into an array of 8×8 elements by making partial saw cuts through the crystal[38]. The cuts are made at various depths, with depth increasing from the centre to the edge of the block. These variations in cut depth result in linear sharing of light among PMTs. Thereby, scintillation light from a photon interaction in the corner element with deep cut will be detected by the PMT located under that element, whereas the light from the middle interaction will be shared by all PMTs[38]. The grooves between the individual elements are filled with an opaque reflective material preventing optical spillover between elements. Modern state of the art scanners employ three to four complete rings of blocks, with about one hundred blocks in each ring and a total axial field of view (FOV) of 15-22 cm[45]. This, depending on the design and purpose of the scanner, may result in as many as 12,000 to 18,000 individual crystal elements used for a single PET system[45]. In another approach, a continuous light guide coupling small crystal elements to an array of photomultiplier tubes is used[36, 39]. For example, a flat panel (9 cm x 18 cm)

containing 4 mm × 6 mm crystal elements (600 crystal elements per panel) is coupled to an array of 15 PMTs. A full ring is assembled using about 28 panels, resulting in 18,000 individual crystal elements per system[45].

In one to one coupling a close-packed array of small discrete detectors is used as the large detector necessary for PET imaging[46]. Direct coupling to discrete crystals and independent parallel signal processing of every electronic channel allows avoiding decoding errors and the identification and potential processing (or rejecting) of multiple interaction events[42]. To improve the spatial resolution in one to one coupling, very small size photo-detectors must be used. This can be achieved either by using APDs or via the coupling of individual channels of a PS-PMT[h] (position-sensitive photomultiplier tube) or MC-PMT[i] (multi-channel photomultiplier tube) to the small crystals[46].

The large number of crystals required, along with a complex assembly and manufacturing process, largely determine the cost of modern PET scanners[45].

2.7 2D and 3D mode of operation

As discussed previously, it is common that modern scanners have at least three complete rings of crystals. Some systems employ movable shielding (septa) between the rows of detectors (crystal rings). These septa are usually made of tungsten (high Z). Deployed septa limit the acceptance angle of LORs by allowing only events from certain rings to be in coincidence with each other. This is known as a two dimensional (2D) mode of operation. In 2D mode, the number of scattered and random events (see section **2.8**) is reduced at the expense of a

---

[h] In PS-PMT, a fine grid dynode structure allows restricting the spread of photoelectrons while in trajectory, thereby making possible a position-sensitive energy measurement within a single photomultiplier tube (PS-PMT) (R. Lecomte, "Novel detector technology for clinical PET," Eur.J.Nucl.Med.Mol.Imaging 36 Suppl 1, S69-85 (2009); P. E. Valk, Positron emission tomography: basic science and clinical practice, edited by P. E. Valk (Springer, London ; New York, 2003), pp. 884).

[i] In MC-PMT technology, the use of a 2D array of glass capillary dynodes, each a few microns wide, has dramatically improved position sensitive energy measurement within PMTs ibid..

reduced efficiency of the PET scanner to true coincidences (see section **2.8**)[45]. In this mode of operation the sensitivity of the PET scanner to true coincidences can be improved by accepting coincidences not only in direct planes, but also in adjacent cross-planes. Direct planes are those planes formed from LORs between detectors within each individual detector ring, while cross-planes are the average of coincidences recorded between detectors in two adjacent detector rings and are the planes falling halfway between direct planes. Acquiring data in this manner improves axial sampling because the events collected by cross planes originate primarily from volumes between the direct planes. Cross-planes together with direct planes produce a dataset consisting of co-planar sets of LORs normal to the axis of the PET scanner. Thus, a system with *p* detector rings will generate a volume of images with *2p-1* planes or slices. For example, a PET camera consisting of 23 rings of detectors spanning the entire axial FOV will produce an image volume consisting of 45 planes or slices.

In a modern scanner it is common for septa to be almost always retracted or even absent altogether. This mode of operation is referred to as three dimensional (3D mode). In 3D mode all coincidences detected within the FOV of the scanner may be accepted. This results in approximately fivefold increase in sensitivity[47]. Due to the increased sensitivity to true coincidences in 3D mode, the same dose to the patient and the same imaging time (as in 2D mode) leads to a substantial improvement in image quality as compared to 2D mode. Also, the increase in acquisition rates for 3D mode allows decreased imaging time (dynamic studies, see section **2.9**) for the same dose and the same statistics, while the same imaging time and same statistics leads to reduced dose (serial studies, see section **2.9**) as compared to 2D mode. The greater sensitivity to true coincidences is achieved at the expense of increased sensitivity to scattered and random coincidences (see section **2.8**). The scatter fraction (ratio of scatter to total) of events detected in a typical thorax PET study in 2D mode is estimated to be about 10 to 20 % and about 40 to 60 % in 3D mode[47].

2.8 PET System Response Components and Necessary Corrections

The events detected in a PET scanner can be one of four types: true, random, scattered and multiple. Typically, 10^6 to 10^9 annihilation events are detected in a typical PET scan[48]. Of these, only true events carry useful spatial information regarding the activity distribution within the patient. Both random and scattered coincidences add a uniform background spanning the imaged volume, hence contributing to image noise and reducing image contrast. Therefore, to minimize the effect of unwanted events and other mechanisms adversely affecting image quality, a number of corrections have to be applied. These corrections include correction for differential detector efficiencies (normalization), random coincidence correction, scatter correction, dead-time correction and attenuation correction.

When a single event with energy higher than the LLD threshold is registered by a detector in coincidence with a valid event in another detector, a coincident event is recorded. The width of the coincidence timing window (CTW) depends on the light decay characteristics of the crystal used and the design of the electronics, while the lower level discriminator (LLD) is set based on the type of crystal and a tradeoff between sensitivity and spatial resolution (typically 350 *keV* for BGO and 420 keV for LSO/LYSO)[45].

Multiple coincidences occur when more than two valid events are detected within the CTW. These types of events are more likely to be observed at high count rates. The ambiguity in assigning a LOR corresponding to a true coincidence correctly results in discarding multiple coincidences under normal circumstances. However, sometimes it may be beneficial to randomly select one of several possible LORs rather than discarding the event altogether[48].

Random coincidences are detected when two independent photons not arising from the same annihilation event are recorded in coincidence. Such a situation is possible due to the finite time resolution of the detectors. A relationship between the random coincidence rate R_{12}, and the single events rates incident upon the relevant detectors R_1 and R_2, respectively can be expressed as:

$$R_{12} = 2\tau R_1 R_2 \qquad \textbf{2.7}$$

where τ is the width of the logic pulse produced when an annihilation photon is absorbed in the detector, otherwise known as the coincidence timing window (CTW). Random coincidences, being relatively uniformly distributed across the FOV, introduce noise into the image causing blurriness and distorted intensity[49]. Random coincidences for each detector pair are usually subtracted by the delayed coincidence method, where the random coincidence rate is estimated during the delayed window[49]. Alternatively, the random coincidence rate in each detector pair can be estimated from the singles rate for each[49].

A *scattered* coincidence is one in which one or both of the detected annihilation photons changed their original flight direction due to Compton interaction within or outside the FOV. If the energy of the scattered photon is greater than the LLD threshold, its detection in coincidence with an unscattered photon will produce a mispositioned LOR. The determining factors for the fraction of scattered events recorded are the size and attenuation characteristics of the object as well as the geometry and energy resolution of the scanner[48]. If not corrected, these scattered events produce a low spatial frequency background that reduces contrast, resulting in a reduced signal to background ratio[48]. In 2D mode, scatter correction is applied using a convolution subtraction method (de-convolution algorithm)[49], where the scatter contribution is reduced via an integral transformation of the projections. Due to the increased sensitivity to scattered coincidences, scatter correction in 3D mode is far more complicated than in 2D mode. A large number of scatter subtraction methods has been explored: convolution subtraction method, Monte-Carlo modeling methods, multiple energy window method, Gaussian fit methods, model-based scatter correction methods and reconstruction-based methods[49].

The majority of single events (about 90% or more) detected by a PET scanner are rejected, because the partner photon has not been registered for various reasons[48]. For example, the second annihilation photon did not deposit sufficient energy in the detector to be recorded or did not interact in crystal at all. Another

reason could be the solid angle coverage around the imaged object by the PET scanner. This type of event contributes to the number of random and multiple coincidences recorded and is a determining factor in problems related to detector dead time[48].

Correction for photon attenuation is by far the largest factor affecting measured counts in PET acquisition. The attenuation of 511 keV annihilation photons within tissue is mostly due to Compton scattering and to a much lesser degree (almost negligibly) photoelectric absorption. The amount of attenuation depends on the electron density of the tissue in question and the total distance traveled by each annihilation photon. To quantify this, if I_0 (counts per second or cps) photons are emitted, then the amount by which the source intensity is reduced is described by:

$$I\left(d_i\right) = I_0 \exp\left(-\mu d_i\right) \qquad \textbf{2.8}$$

where d_i is the distance traveled by the annihilation photon prior to interaction in the i^{th} detector and μ (cm^{-1}) is the linear attenuation coefficient. The linear attenuation coefficient is a function of the physical properties of the attenuating tissue (atomic number and density) and the energy of the annihilation photons. It is defined as the product of interaction cross section and number of targets per unit volume. For 511 keV photons in water, μ_{water}(@ 511 keV) = 0.09695 cm^{-1}, which gives a half-depth, the depth at which 1/2 the initial photons have been absorbed, of 7.15 cm.

Assuming an organ or tissue of uniform density, the probability of detecting both annihilation photons in coincidence is given by:

$$P = \exp\left(-\mu d_1\right) \cdot \exp\left(-\mu d_2\right) = \exp\left(-\mu\left[d_1 + d_2\right]\right) = \exp\left(-\mu D\right) \qquad \textbf{2.9}$$

where μ is the linear attenuation coefficient, $\exp(-\mu d_1)$ is the probability of detecting photon 1, $\exp(-\mu d_2)$ is the probability of detecting photon 2, and D is the sum of tissue thicknesses (d_1 and d_2) traversed. This probability represents the *attenuation correction factor* (ACF) for a given tissue and is independent of the depth of the annihilation event in the patient and is the same as for an external

source located along the LOR. Thus, a theoretical calculation of ACFs may be done based on equation (**2.9**). This method (also referred to as Chang's method) assumes prior knowledge of μ, the shape of the organs involved and uniform attenuation (for example in the head)[38]. It is not well suited for ACF calculations in situations where imaging of the thorax and the abdomen is involved, due to the non-uniform attenuation resulting from the prevalence of various tissue structures[38]. Alternatively, a thin rod source is used to perform a transmission scan which allows the measurement of regional attenuation correction factors[36]. This source is attached to the PET scanner gantry and contains a relatively long-lived positron emitter. The ACFs are calculated for each LOR based on a blank scan without the patient in the scanner and a transmission scan with the patient positioned in the scanner as:

$$ACF_{ij} = \frac{B_{ij}}{R_{ij}} = \exp\left(\sum_{k=0}^{n} \mu_k D_k\right), \quad i, j \Rightarrow LOR \qquad \textbf{2.10}$$

where ACF_{ij} refers to the attenuation correction factor for detector pairs i and j (i.e., LOR), B_{ij} and R_{ij} refer to the counts measured by detector pairs i and j in the blank and transmission scans respectively, μ_k and D_k are the linear attenuation coefficient and thickness of the kth organ or tissue, and n is the number of organs or tissues that the photon travels through. Some sort of noise suppression, such as a space-invariant filter[50], is applied to the blank and transmission data sets prior to ACF calculations to alleviate the effects of random coincidence subtraction (negative and zero values) and other factors contributing to image noise discussed previously. In 3D mode, attenuation correction is more complicated than in 2D mode due to an increased scatter fraction and excessively long dead time experienced by detectors closer to the source. This can be dealt with by using 2D transmission data for correction of 3D emission data. Another approach is to use a much weaker point source for performing transmission scans [49]. Alternatively, a well collimated single photon emitter producing photons with energies close to *511 keV* is used. A common radionuclide used for the transmission source is ^{68}Ge, which decays with a half-life of 287 days to the positron emitting isotope ^{68}Ga. Due to the relatively short half-life of ^{68}Ge, ^{137}Cs (single photon emitter @ *662*

keV) is becoming the transmission source of choice in standalone PET systems because of its long half-life of 30 years. The calculation of ACFs is much simplified in PET/CT scanners, where CT transmission scans, in addition to their use in fusion of PET and CT images, are used to obtain ACFs at 511 keV. Because CT transmission scans are normally acquired at about 70 keV and ACFs depend on photon energy, the data derived must be scaled to the 511 keV photons of the PET scanner. CT values represented in Hounsfield units are mapped to the ACFs by means of methods such as bilinear mapping[41]. The advantages of the CT transmission method over the transmission method using radioactive sources are total duration of the scan and essentially noiseless attenuation correction factors[38]. The duration of a typical CT scan is a minute at most, while a ^{68}Ge transmission scan can take as much as 20-40 min depending on the activity level of the source[38].

Dead-time is defined as the time interval during which a detector, having just responded to one event, is unable to respond to another. It is related to the signal integration time that depends on electronics and scintillation decay time[38]. It can seriously affect measured data at high count rates (especially transmission data used for ACF corrections), producing artifacts (structured noise) and contributing to noise in reconstructed images[49]. One way to reduce dead-time is to use detectors with shorter scintillation decay time, high-speed electronics and pulse pile-up rejection circuits[38]. Practically, an experimental dead-time model is used to produce *dead-time correction factors* from observed count rates as a function of increasing activity concentrations. A dead time curve is calculated from a series of measurements done on a uniform phantom, starting with high radiotracer concentration down to low concentration until dead-time losses are negligible[49].

Due to differences in the physical properties of the individual detectors, as well as detector efficiencies, normalization is performed to equalize or make uniform (as nearly as possible) responses of the detectors to the number of incident radiation events in the FOV. The detector efficiency correction is necessary as crystals positioned in the center of the detector block are more efficient than those at the edge of the block[49]. Correction factors are obtained

using an external positron source. A lengthy (*~12h*) scan using an external uniform source is acquired such that all LORs are exposed to approximately the same coincident photon flux and the correction factors are calculated as[38, 41]:

$$F_{ij}^{norm} = \frac{N_{ij}}{\frac{1}{m}\sum_{l=1}^{m} N_{ij}^{l}} \qquad \textbf{2.11}$$

where F_{ij}^{norm} refers to the *normalization correction factor* for detector pairs *i* and *j* (i.e., LOR), N_{ij} is the number of counts in the LOR and *m* is the number of LORs. It should be noted that same-ring coincident data and adjacent-ring coincidence data must be normalized separately in 2D mode as the two geometries have different detection efficiencies.

The total number of prompt coincidence events, P_{ij}, measured by detector pairs *i* and *j* can be expressed as:

$$P_{ij} = F_{ij}^{norm} \cdot \left(\frac{T_{ij}}{ACF_{ij}} + S_{ij} + R_{ij} \right) \qquad \textbf{2.12}$$

where F_{ij}^{norm} refers to the normalization correction factor for detector pairs *i* and *j* (i.e., LOR); ACF_{ij} is the attenuation correction factor; T_{ij}, S_{ij} and R_{ij} refer to the true, scattered and random coincidences measured by detector pairs *i* and *j* respectively.

2.9 Scan Modes (static or dynamic)

There are two scan modes in PET: static and dynamic. These should not be confused with the 2D and 3D modes of operation discussed in section **2.7**. In static imaging (nowadays 3D PET), a static image is acquired over the total length of the scan, whereas in dynamic imaging (4D PET), data is collected in multiple frames, each of a predetermined duration. In dynamic imaging, a gated method can be employed to reduce artifacts induced by respiratory motion. Whereas static scans are useful to estimate gross tracer uptake, dynamic scans reveal the time course of activity distribution in tissue[38]. Depending on the scan mode used and the reconstruction algorithm employed, the annihilation events in PET are

generally recorded either in list or histogram mode. In list mode, each coincidence event meeting both energy and timing criteria is individually written to a file by sorting hardware, with information about the two locations at which the annihilation photons interacted and the time at which the event occurred[48]. In histogram mode, a memory location assigned to each possible LOR is incremented by *1* for every valid event detected in that LOR, thus providing the integrated number of events detected in each LOR[48]. With the exception of short duration acquisitions on scanners with a very large numbers of LORs, where the average number of events per LOR is less than one, the histogram mode is often the most efficient way of storing data. In histogram mode, the time interval over which the events are integrated has to be specified prior to data acquisition, making this mode suitable for static scans. With the introduction of faster computers, the list mode acquisition is becoming more commonly used in modern scanners. Also, list mode acquisition is preferred in dynamic studies since the data can be sorted into time bins (frames) after the completion of the study, providing variable temporal resolution[38, 48].

2.10 Image Reconstruction

As previously mentioned, in PET imaging, recorded events represent the summation of activity within the tissue along a LOR. The data is stored in sinograms[j], where each LOR (and hence, detector pair) corresponds to a particular pixel (or element) in the sinogram, characterized by two orthogonal coordinates[38]. In other words, each LOR is defined by its distance (r) from the center of the gantry and the angle of orientation (φ) of the LOR (i.e., the angle between r and the vertical axis of the FOV)[38]. In order to reconstruct a volumetric

[j] Typically, for detector ring with N detectors, each detector can be in coincidence with up to $N/2$ detectors on the opposite side, resulting in $N/2$ fan-beam projections for each detector. These projections form an angle of acceptance for each detector in transaxial plane, while these angles for all detectors form the transaxial FOV. A sinogram is essentially a matrix of fan-beam projections defined by the angle of acceptance (G. B. Saha, Basics of PET imaging: physics, chemistry, and regulations, edited by G. B. Saha (Springer, New York, NY, 2005), pp. 206).

image representing the distribution of activity, large number LORs are collected and, after all corrections (scatter, randoms, attenuation, etc.) have been applied to the data, are processed according to an image reconstruction algorithm. A wide variety of PET reconstruction algorithms exists, owing to the absence of a uniquely optimal technique. These algorithms can be broadly divided into two categories: analytical approaches (based on the Fourier transform, a linear operations) and iterative methods (sometimes called Non-Fourier-based or non-linear operations)[49]. The most common algorithm used for data acquired in 2D mode has been filtered back-projection (FBP) from fan-beam projections. 3D reconstruction algorithms are much more complex and a wide variety of cone-beam (3D FBP) and iterative techniques have been investigated. 3D reconstruction techniques are very computing intensive (both in terms of time and memory required) due to the enormous number of LORs acquired. Often it is more advantageous to reduce the redundant 3D data set to synthetic 2D data which can be processed by analytic or iterative 2D algorithms. To this end, different hybrid algorithms combining fast Fourier rebinning (FORE, FOREX, FORE-J) with iterative 2D algorithms have been introduced[39]. 3D Re-Projection (3D-RP)[49, 51] is another approach based on restoration of the missing elements of projections. This is done through numerical forward projection from an initial estimate of an image formed by reconstructing (2D FBP) only direct planes.

The back-projection technique, a workhorse of tomography, has been around since the early 1960s and involves direct re-projection of multiple projections of 2D activity distribution to a common image plane[49]. A variation of this method involving filtering of the observed projections prior to back-projection has been proposed and implemented by Chesler in 1972, in order to reduce blurriness characteristic to images reconstructed using the original algorithm[49]. A simple reconstruction procedure using FBP of data acquired in 2D mode involves convolution of the Fourier transformed (1D transform) angular projection data with a ramp filter in the frequency domain prior to back-projection to a common image plane using the inverse Fourier transform. Use of the ramp filter is necessary to reduce the structured noise present in PET data. However, while

reducing structured noise during the back-projection process, the ramp filter also has the property of amplifying high-frequency components. Since statistical noise characteristic to PET data manifests itself in Fourier space as high-frequency components, the process of FBP amplifies statistical image noise. For this reason, modifications of the ramp filter which discard frequencies higher than a certain limit are normally used. Perhaps the most common of these is the Hann filter (also known as an apodizing or tapering filter). Such low-pass filters are equivalent to some form of smoothing in image space. The traditional Hann window cutoff is at the Nyquist frequency (half the sampling rate).

Of all iterative algorithms introduced between 1970 – 1980s, the most widely used in PET are the ML-EM (Maximum-Likelihood Expectation Maximization) algorithm and its accelerated version OSEM (Ordered Subset Expectation Maximization)[39, 49]. With its roots in astronomy (Richardson-Lucy algorithm), ML-EM was first introduced by Dempster and Laird in 1977 and first applied in PET by Shepp and Vardi and Lange and Carson in the early 1980s[39]. Iterative algorithms offer less streak artifacts and lower noise levels compared to FBP. However, up until 1994 when the OSEM algorithm was introduced, iterative techniques were not sufficiently fast for clinical applications and FBP was the method of choice in most commercial systems. Since then, the OSEM and its variants have become very popular. The basic idea behind iterative reconstruction approaches is maximization or minimization of a cost (also known as target or objective function) function depending on the algorithm used. In the first step an initial guess is made of the image intensity distribution (often a blank or uniform grayscale image). The next step is to forward-project it into the projection domain, i.e. sum up all the activity in pixels that are intersected by the line of response that corresponds to the measured sinogram element[48]. Once this process is complete, the projections obtained from the initial image (guess) are compared with the measured projections. Based on the differences between observed and measured projections, correction factors for all projections are generated. The initial guess is then adjusted by multiplying it with these correction factors back-projected into the spatial domain (image domain) and dividing it by "weighting

terms" based on the acquisition system model, thus generating a new corrected image[49]. The whole process is repeated iteratively using the newly corrected image as the initial image until a certain likelihood criterion is reached[48, 49]. Iterative algorithms differ by their cost function, update (or search) function, and the model describing the acquisition system (including noise)[1]. The cost function gives a measure of the difference (or similarity) between the image estimate and actual measured data, while the update function is an algorithm that determines the next estimate of the image. An additional advantage of iterative algorithms over FBP is the possibility to include prior information about the object scanned into the reconstruction process. Examples of such prior information include level of smoothness and non-negativity constraints. Both ML-EM and OSEM approaches maximize Poisson likelihood. The main disadvantage of ML-EM is its speed (low convergence rate) and uncertainty in determining the number of iterations required for obtaining acceptable results. An excessive number of iterations in ML-EM leads to image quality degradation due to image noise amplification[49]. In order to overcome long processing times typical of ML-EM, in the OSEM algorithm projections are grouped into a number of subsets in which projections are uniformly distributed. A standard EM algorithm is applied to one of the subsets and the result becomes the starting value to be used with the next subset. In a single OSEM iteration this process is repeated until each subset has been used once[1]. The cost function is updated as many times as the number of subsets, proportionally accelerating convergence. A popular variation of OSEM is Attenuation-Weighted OSEM (AW-OSEM)[49]. Two other examples of iterative algorithms applied to PET are Rescaled Block-Iterative ML-EM and Row-Action Maximum-Likelihood Algorithm (RAMLA)[39, 49]. In block-iterative (or raw-action) algorithms only a subset of the data is used per iteration. RAMLA is a special case of OSEM requiring sequences of orthogonal projections. This leads to faster convergence than OSEM itself[38]. These algorithms are akin to OSEM, but guarantee an asymptotic convergence of Poisson likelihood estimation (maximization) under certain conditions[39]. Practically all commercially available scanners allow for both FPB and iterative reconstructions, both with and without

attenuation correction. Iterative techniques such as the OS-EM and RAMLA (2D or 3D) algorithms have almost completely replaced FPB in clinical applications.

2.11 Absolute Quantification

Absolute quantification (AQ) is used to quantify patient images in order to distinguish benign tissue from malignant and (or) to assess response to therapy. AQ is a term referring to the process of expressing the activity distribution in a scanned object in terms of activity per unit volume (activity concentration). This is achieved by applying a calibration factor to voxels of a reconstructed tomographic image that maps counts per second (cps) into activity per unit volume (mCi/cc or MBq/cc). The calibration factor is derived from the volume image obtained by scanning a uniform cylindrical water phantom of known volume, injected with a known amount of activity[41] (i.e. known activity concentration). From reconstructed images, the average count density, C (counts per voxel per second), within the central portion of the phantom image is determined and a calibration factor (*CF*) relating image counts and known activity concentration, *A* (μCi/cc), is deduced according to[48]:

$$CF = \frac{A\left(\mu Ci/cc\right) \times BF}{C\left(counts/voxel/\sec\right)} \qquad \textbf{2.13}$$

where *BF* is the branching fraction. *BF* represents the fraction of decays that occur via positron emission, since the radionuclides used in PET do not necessarily decay *100%* by positron emission (see **Table 6.2**).

The uptake of a radioactive tracer in a given volume of interest (VOI or in the 2D case ROI – region of interest) can be normalized by correcting for differences in total injected dose and body weight. The normalized uptake, in terms of activity concentration, is a semiquantitative determination of tissue activity. It is called the standardized uptake value (SUV) and is defined as:

$$SUV\left(g/cc\right)=\frac{Tissue\ Activity\ Concentration\left(\mu Ci/cc\right)}{Total\ Injected\ Dose\left(\mu Ci\right)/Body\ Weight\left(g\right)} \qquad \textbf{2.14}$$

Under some circumstances it is more convenient to express SUV as a unitless parameter after dividing it by the density of tissue which is often assumed to be ~ 1 g/cc (or g/cm^3)[52, 53]. A unitless SUV can be expressed in terms of annihilation event counts rather than activity concentration.

One of the advantages of PET over some other imaging modalities is that it can accurately determine the activity concentration of radiotracer within a volume, allowing (for example) the classification of a lesion in terms of its metabolic rate[38]. FDG (^{18}F-fluoro-deoxy-glucose) has been shown to be very effective in characterizing tumors by measuring glucose metabolism. Oncology scans using FDG make up over 90% of all PET scans in current practice. A typical dose of FDG used in an oncological scan is 200-400 MBq for an adult human. This tracer is glucose analog and follows a similar metabolic pathway. Because the oxygen atom is replaced by ^{18}F to generate FDG, it is not metabolized to CO_2 and water and remains trapped within tissue. The result of this is intense radiolabeling of tissues with high glucose uptake, such as the brain, the liver, and most cancers. If a static sequence of images has been acquired, images show FDG-6-PO_4 within tissue and FDG in the tissue and plasma. There are a number of factors affecting the amount of activity (SUV) measured using static imaging: blood glucose, partial volume effect, body composition and habitus (e.g., fat has lower uptake), and the time elapsed between tracer injection and the start of the PET scan (uptake period)[41]. These factors lead to tracer distribution (SUV) by static imaging being regarded as a *semiquantitative measure*[41]. In a dynamic scan, the ROIs or VOI defined by the operator can be applied to the same region (or volume) on all images to generate a time activity curve (TAC). TACs express the activity concentration in a specific region of the body over time[48]. The TACs can then be used in the framework of a compartmental model[1, 39] to determine biologically meaningful parameters (metabolic rate for instance) and to construct parametric images[48]. The metabolic rate obtained from kinetic modeling is also referred to as a *semiquantitative*

measure[41]. The most significant advantages of dynamic imaging over static imaging are the ability to quantify the actual glucose metabolism and the avoidance of the aforementioned factors affecting the activity distribution measured via static scans[41].

3. Properties of Noise in PET Images Reconstructed with FBP and RAMLA algorithms[k]

One of the confounding factors in the interpretation of PET images is image noise. Clinical PET images typically display increased noise levels compared to other modalities such as CT and MRI. Effective image noise reduction is greatly dependent on an accurate knowledge of the parameters which characterize this noise. Unfortunately, detailed properties of the noise affecting clinical PET images are not often well characterized. This chapter examines the statistical properties of noise in PET images acquired with a GEMINI PET/CT scanner (Philips Medical Systems Inc.).

3.1 Sources of Noise in PET

In this section we summarize the most important sources of error manifesting as noise in reconstructed PET images that cause them to differ from the actual radiotracer distribution within the body. Noise observed in PET images leads to loss of spatial and temporal resolution; lower acquired signal level; missing features, which are present in the body but do not appear in the images; and false artifacts and textures, which appear as features in the image, but are not actually present in the body[54].

The preceding chapter regarding positron physics and PET technology has identified several phenomena that degrade the reconstructed image quality, namely:

Poisson counting noise, radioactive decay obeys Poisson statistics

Positron range, leading to loss of resolution

[k] A version of this chapter has been published: A. Teymurazyan, T. Riauka, H. -S Jans, et al., "Properties of Noise in Positron Emission Tomography Images Reconstructed with Filtered-Backprojection and Row-Action Maximum Likelihood Algorithm," J.Digital Imaging, 1-12 (2012).

Non-colinearity effect (annihilation photons are not always emitted at 180° from each other), the assumption of colinearity in the reconstruction algorithms leading to loss of resolution

Radial blurring due to detector dimensions (use of thin, long detector crystals), leading to loss of resolution

Acquisition Mode (3D or 2D), higher statistical noise in 3D mode

Blank, Transmission and Emission scan, contributing to statistical noise

Scan duration, longer scan times lead to lower relative noise

Out of field counts, contributing to false coincidences

Multiple coincidences, leading to counts at erroneous positions

Random coincidences, leading to counts at erroneous positions

Scattered coincidences, leading to counts at erroneous positions

Photon attenuation, leading to an underestimation of activity inside the body

Detector dead time, leading to loss of counts

Normalization (detector efficiency correction), despite the best efforts, differences in detector efficiencies may contribute to structured noise in the acquired data

In addition to these, patient and organ motion during scan time and image reconstruction artifacts can also introduce additional noise into reconstructed PET images. The magnitude of the noise in reconstructed PET images is affected by: the PET scanner used; acquisition mode (2D or 3D) used; scan duration; amount of administered tracer; geometry of tracer distribution; applied correction methods; radioactive decay; and reconstruction algorithm (analytical or iterative)[49].

The statistical properties of the noise in PET images are discussed in the following sections.

3.2 Statistical Properties of Noise in PET Imaging

Positron emission itself is well characterized by a Poisson distribution[55-57]. The scanner's detection system and other electronic components then add their own characteristic noise to this initial Poisson distribution. The resulting noise distribution is further altered by corrections processing and image reconstruction[58, 59]. Some reconstruction schemes, such as the EM-ML[60], are explicitly based on the assumption of Poisson statistics in the acquired sinograms. This assumption, given all the pre-processing steps prior to reconstruction, may prove to be invalid. Filtered-Backprojection[61], on the other hand, is predicated on the analytic inversion of noise free projection data. The degree to which the Poisson characteristics of PET noise are preserved is highly dependent on the manner in which the raw data is processed[62]. Noise propagation is affected by machine type, acquisition mode, scan time, amount and distribution of tracer, applied corrections, and the reconstruction algorithm[1, 49, 54]. For commercial systems, the details of clinical image production are usually held proprietary by the vendor. This effective black box nature of the process necessitates an empirical evaluation in order to characterize the noise present in images presented to the user.

Post-processing may be employed to reduce the level of noise present in clinical PET images. The effectiveness of post-processing is greatly aided by a proper quantification of the characteristics of the noise present. An assumption common to many post-processing approaches is that overall the noise may be characterized as Gaussian[63, 64]. Alternatively, the noise in clinical PET images has been described using a Poisson + Gaussian model[65], where the presence of both, correlated and uncorrelated components, is assumed. Other studies[66-69] are aimed specifically at the reduction of Poisson noise contained in medical (including PET) images, exploiting its statistical properties. It is not clear if this approach is strictly valid in the case of PET images[70].

In this chapter the statistical properties of PET images acquired according to clinical protocol and reconstructed with Filtered-Backprojection[61] and Row-Action Maximum Likelihood Algorithm[71] are examined. Filtered-Backprojection

(FBP) has significant limitations compared to more general maximum likelihood (ML) based iterative reconstruction methods. FBP does not take into account counting statistics, assumes shift invariance, treats LORs (lines of response) as close approximations to line integrals, and is often limited to approximate empirical scattering corrections. Iterative reconstruction methods, in contrast, do not rely on assumptions of well-behaved LORs and shift invariance, and uniform sampling is not necessary. Factors like detector resolution, scattering, attenuation, positron range, photon non-colinearity can be explicitly incorporated into the probabilistic calculation of positron annihilation detection along a particular LOR[72]. Several groups[73-75] have performed theoretical analyses of the noise properties of images reconstructed with both methods. FBP tends to spread noise variance from high count regions to low count regions, producing increased noise correlation with decreasing FBP filter cutoff frequency. This results in a more uniform noise variance[54]. RAMLA, which was proposed as a faster alternative to the EM algorithm[76] and can be considered as a special case of the Ordered Subsets Expectation Maximization algorithm[75, 77], yields significantly decreased noise variance in low count regions compared to FBP.

3.3 Experimental Evaluation of the Statistical Characteristics of Noise in Reconstructed PET Images

3.3.1 Methods and Materials

Statistical properties of noise were evaluated for PET images acquired with a Philips Gemini GS PET/CT scanner (produced in 2003). It is an integrated system that consists of an Mx8000 Dual-Exp CT system for CT imaging and an Allegro PET system for PET imaging. In addition to its CT based attenuation correction capability, this first generation Gemini system inherits the transmission scan mechanism of the Philips Allegro system that uses a singles transmission source (Cs-137). The PET scanner is comprised of 28 flat modules of a 22×29 (tangential and axial directions) array of GSO crystals, which form 29 rings with 616 crystals per ring. The dimensions of the crystals are 4×6×20 mm^3 in the tangential, axial,

and radial directions, respectively. Data acquisition was performed in list mode. The data acquired from the scanner are binned into a sinogram with 161 angles and 295 rays for every ring combination (total of 29^2 = 841 combinations). Interpolation is performed to rebin these data into a 256×192 (rays × angles) sinogram and 7 tilts (out-of-plane angle). The whole-body field of view (FOV) for Gemini PET is 576 mm transaxially and 183 mm axially. **Table 3.1** lists the characteristics of Gemini PET scanner. The noise measurements were performed with a cylindrical phantom (long axis coincident with the reconstruction centre and orthogonal to the image plane), which was 20 cm in diameter and 35 cm in length.

Table 3.1: Philips Gemini GS PET/CT characteristics (produced in 2003)

Parameter	Specifications
Number of blocks	28
Number of Detector rings	29
Maximum ring difference	28
Number of crystals	17864 GSO
Number of PMT's	420
PMT diameter	39 mm
Crystal dimensions	4 (transaxial)×6 (axial)×20 (radial) mm^3
Detector Ring diameter	800 mm
Patient portal diameter	565 mm
Axial FOV	183 mm
Number of Image Planes	90 or 45 (brain and whole body, respectively)
Plane Spacing	2 or 4 mm (brain and whole body, respectively)
Transmission source	Rotating 740 MBq ^{137}Cs point source
Reconstruction Algorithms	FORE3D+ FBP
	RAMLA3D

The phantom was injected with 87 MBq of ^{11}C ($T_{1/2}$ = 20 min) and well mixed to provide a uniform activity distribution. The phantom was scanned in a single bed position. A dynamic sequence of 20 frames was acquired for 100 min according to the following schedule: 20×300 s. For each frame the reconstructed image size was 144×144 pixels in 45 slices (144×144×45), with a pixel size of 4×4 mm^2 and slice thickness of 4 mm. Statistical properties of noise were evaluated for PET images reconstructed according to current clinical settings used at the Cross Cancer Institute. Routine clinical image reconstruction is performed with a fast, fully 3D iterative algorithm (3D-RAMLA) with two iterations, relaxation parameter of 0.006 and a "blob" radius of 2.5 pixels. For comparison, statistical properties of noise were also evaluated for images reconstructed with Fourier rebinning (FORE) followed by a Hanning filtered-back-projection (FBP) algorithm. A default value of 3.00 for the Hanning smoothing parameter, supplied by the scanner manufacturer, was used in reconstruction. PET scans were normalized to correct for the variation in detector efficiencies and distortion. Emission data were corrected for randoms, scatter, attenuation and decay. The randoms correction was accomplished via direct online randoms subtraction from the prompt sinograms (direct randoms estimation using a delayed coincidence window technique, no smoothing of randoms). The decay correction was performed to the beginning of the scan. For RAMLA reconstruction the attenuation map was obtained from the CT scan and scatter correction was accomplished by a single-scatter simulation (SSS) technique (the scatter sinogram is subtracted from the non-scatter corrected sinogram). In the Gemini system, CT-based three-dimensional attenuation correction (CT-3DAC) is incorporated into the three-dimensional row-action maximum likelihood algorithm (3D-RAMLA) for PET image reconstruction. For data reconstructed with FBP the scatter correction was accomplished by a uniform background subtraction (UNI-BGSUB) technique, while attenuation correction was performed by the reconstruction-reprojection method based on the reconstruction and forward projection of transmission images. 3D-RAMLA and Hanning FBP reconstructions were performed using default (scanner manufacturer supplied) reconstruction

tools. The software used in the analysis of the reconstructed image data was developed in MATLAB 7.4.0, on a PC.

The noise probability density function (PDF) may be characterized by examining the histogram of a region of interest (ROI) selected in a uniform image[78]. A circular ROI covering the inner 79% (1542 pixels, where each pixel is 4 by 4 mm^2) of the phantom's cross section was selected from each slice. The outer edge of the ROI was kept inside the full active extent of the phantom in order to avoid confounding partial volume effects which occur in the immediate vicinity of the phantom wall and to avoid the air bubble visible on **Figure 3.1**. In the longitudinal direction the analysis spanned the entire 45 slices acquired in a single bed position. Histograms generated from these ROI's were fitted, using maximum likelihood estimation, to Poisson, Normal (Gaussian), Negative Binomial, Log-Normal and Gamma distributions. Derived in this manner, the parameters specific to each distribution model are then used to calculate Cumulative Distribution Functions (CDFs), Quantiles (inverse of the CDF), Skewness and Kurtosis (excess kurtosis).

Several graphical methods may be used to evaluate the differences between the probability distribution of a population from which a random sample is drawn and that of a reference distribution. Here, we employ commonly used Quantile – Quantile (Q-Q) plots[79] for this purpose. In a Q-Q plot, the inverse of the cumulative distribution function (iCDF) of experimental data (experimental Quantiles) is plotted against the iCDF of the distribution fitted to the data (theoretical Quantiles). If the experimental distribution is the same as the reference distribution then the resulting Q-Q plot will follow a 45° line rising from left to right. Linear plots which deviate from the 45° line indicate a difference in dispersion between the two distributions. Substantial deviations from linearity dictate a rejection of the hypothesis of sameness. Quantiles of the experimental PET distribution were compared to Quantiles of the Poisson, Negative Binomial, Normal, Log-Normal and Gamma distributions. The Shapiro-Wilk, Anderson-Darling, Kolmogorov-Smirnov, Pearson's Chi-square, D'Agostino's K-squared, and a score of other tests may also be used to evaluate

the distribution of noise. Quantile-Quantile plots were chosen for this investigation because, unlike some of the tests mentioned above, Q-Q plots: (1) can be applied to both continuous and discrete distributions, (2) do not require sample sizes to be the same, and (3) allow testing of many distributional aspects such as shifts in location, shifts in scale, changes in symmetry and the presence of outliers simultaneously[79]. A pseudo-Poisson model has been proposed in the literature[58] for simulating PET noise. Using this model to simulate an over-dispersed Poisson distribution and estimate its variance, the mean number of counts is scaled by an empirically determined parameter. The same can be achieved by means of a Negative Binomial distribution without introducing scanner dependent empirical relationships. The Negative Binomial distribution was included in this comparison as it is the simplest way of modeling Poisson distributions in application to situations in which the variance is higher than the mean number of counts (over-dispersed Poisson distribution). This situation can be viewed as a Poisson model with Gamma heterogeneity, where the Gamma noise has a mean of one. In other words, the Negative Binomial distribution can be described as a continuous mixture of Poisson distributions, while the mixing rate is characterized by a Gamma distribution and accounts for over-dispersed or correlated Poisson counts[80]. The Gamma distribution was considered because it describes Poisson processes and, with an appropriate choice of shape and scale parameters, is capable of modeling an over-dispersed Poisson process. Lastly, it has been suggested in the literature that the histogram of a ROI selected in a uniform PET image may be well approximated by Log-Normal distribution[81] and thus this distribution was also investigated.

Skewness (third standardized moment) is a measure of the asymmetry of a probability distribution and is mathematically defined in terms of the third moment about the mean and standard deviation as:

$$Skewness = E\left[\left(\frac{X - \mu}{\sigma}\right)^3\right] \qquad \textbf{3.1}$$

where X is the random variable, μ is the mean, and σ is the standard deviation and E is the expectation operator. Kurtosis (fourth standardized moment) is a measure

of the "peakedness" of a probability distribution. The greater the Kurtosis value, the greater the contribution to the variance from large deviations. Conversely, a variance composed of modest deviations results in a smaller value of Kurtosis. Thus a high Kurtosis value corresponds to a sharply peaked distribution with long ample tails while a low Kurtosis is descriptive of a more rounded peak with shorter thinner tails. Mathematically, Kurtosis is defined in terms of the fourth moment about the mean and standard deviation as:

$$Kurtosis = E\left[\frac{\left(X - \mu\right)^4}{\sigma^4}\right] \qquad \textbf{3.2}$$

In this work, Skewness and Excess Kurtosis (Kurtosis – 3) are used in conjunction with CDF's and Q-Q plots to examine the statistical properties of noise in PET data. These properties are used to evaluate the experimental distribution in comparison to five model distributions (Poisson, Normal, Negative Binomial, Log-Normal and Gamma). The performance of the Normal, Negative Binomial, Log-Normal and Gamma distributions is compared by means of Root Mean Square Error (RMSE), calculated for Normalized Standard Deviation (NSD), Skewness and Excess Kurtosis. In the 2D case, NSD was defined as the standard deviation (STD) scaled to the Mean of that slice, while in the 3D case it was defined as standard deviation (STD) of the phantom volume of interest (VOI) scaled to its Mean. Analysis of goodness-of-fit between the experimental distribution and he five model distributions is performed by means of Q-Q plots for both 2D and 3D case. Kurtosis and Skewness are utilized to provide an overall feel for the shape of the experimental data and a further measure of the fit to these five distributions.

The spatial variation of noise within the phantom was investigated along single pixel width vertical and horizontal diametric profiles. The data obtained from these profiles is suggestive of certain trends but plagued by excessive noise. Five circular sub-regions (top, bottom, left, right, and center) within the area of interest were also analyzed. These procedures were applied to images

reconstructed from the same raw data set with both Filtered-Backprojection (FBP) and the Row-Action Maximum Likelihood Algorithm (RAMLA).

3.3.2 Results and Discussion

The simple case of a circular ROI in the uniform phantom was investigated first. **Figure 3.1** shows images from slice 21 (first time frame) located near the centre of this uniform phantom reconstructed using RAMLA and FBP and their respective ROI histograms.

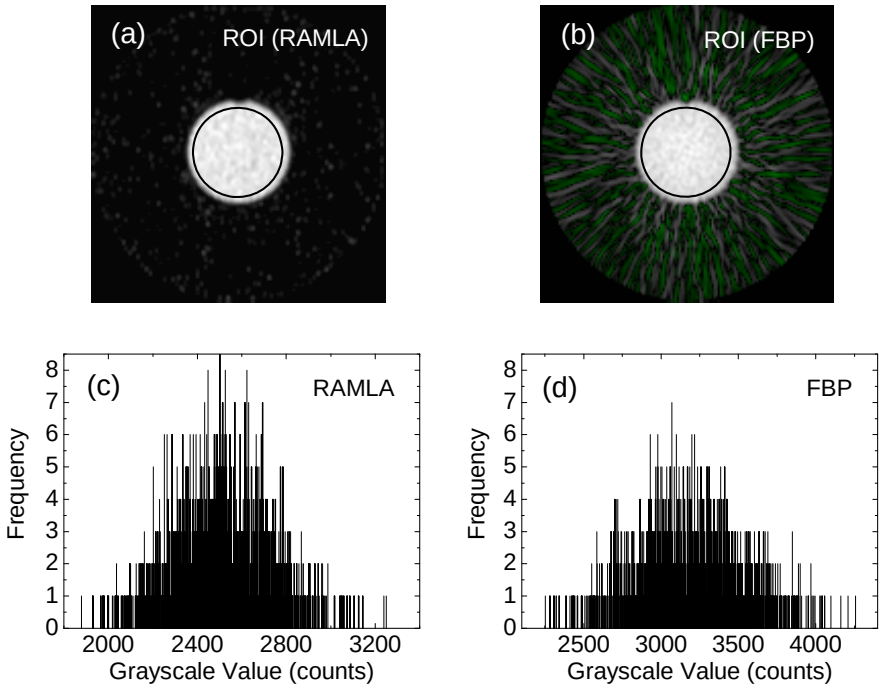

Figure 3.1: 4D PET study on a cylindrical phantom (1st time frame of slice 21). ROI was selected on image reconstructed with RAMLA (a) and FBP (b). The histogram generated from ROI selected on image reconstructed with RAMLA (c) and FBP (d).

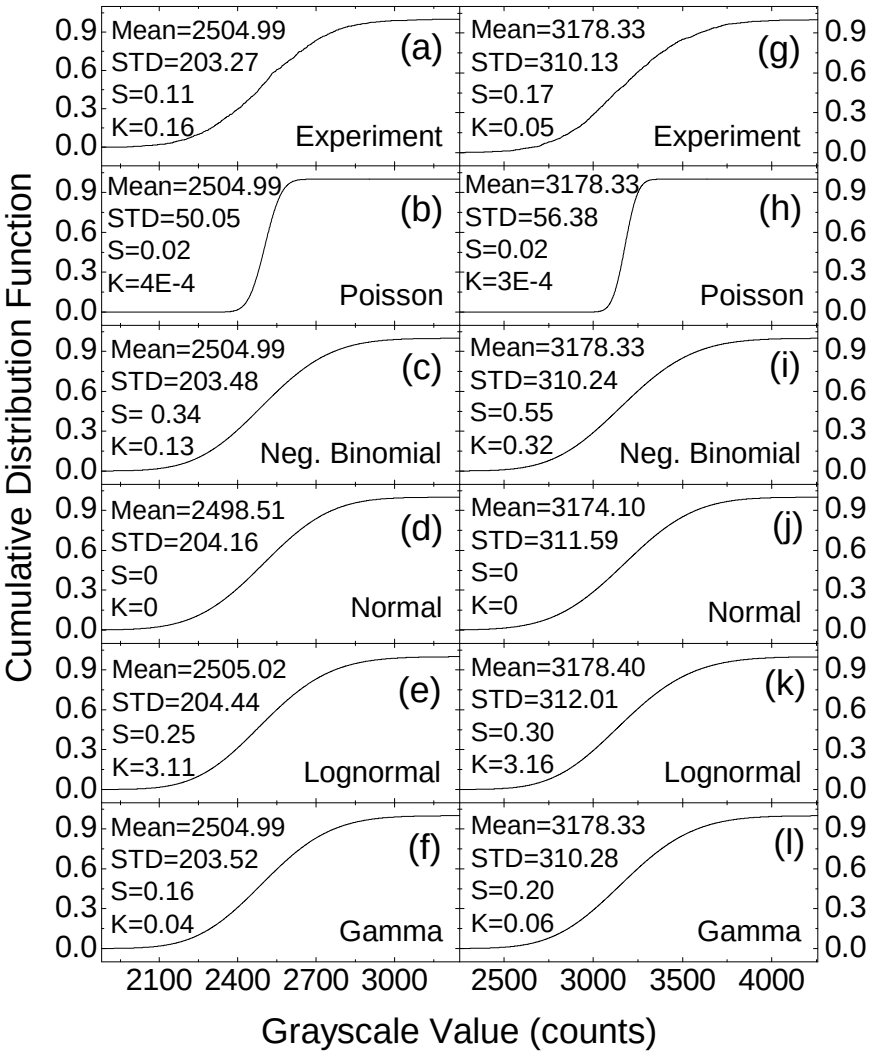

Figure 3.2: Graph of the CDFs plotted vs. pixel counts. Experimental CDF is compared to CDFs of Poisson, Negative Binomial, Normal, Log-Normal and Gamma distributions. The theoretical CDFs are calculated using maximum likelihood estimates for the parameters. The histogram (1st time frame of slice 21) for the fits was generated from a circular ROI selected on a 2D image reconstructed with RAMLA (a through f) and FBP (g through l).

Immediately obvious are the higher reconstructed counts seen in **Figure 3.1** produced by FBP as compared to RAMLA from the same raw data. This can be attributed to a difference in the preprocessing steps (scatter and attenuation correction) and reconstruction (full 3D-RAMLA vs. FORE+FBP). Based on the process of maximum likelihood estimation (MLE), Cumulative Distribution functions (CDFs), Quantiles, Skewness, and Excess Kurtosis were calculated in order to model the experimental data with Poisson, Normal, Negative Binomial, Log-Normal, and Gamma distributions.

Cumulative Distribution functions (CDFs) and cumulants of these distributions calculated from maximum likelihood estimates (MLE) of the distribution parameters are presented in **Figure 3.2** and **Table 3.2** along with respective Margins of Error for the 95% confidence level for slice 21 (1^{st} time frame). Margins of error are calculated as the half widths of the 95% confidence intervals for the parameter estimates. Exploring the representative data of slice 21 (1^{st} time frame), it is readily evident from **Figure 3.2** and **Table 3.2** that the noise characterizing this ROI is not Poisson distributed but is instead better modeled by the Negative Binomial, Normal, Log-Normal and Gamma distributions.

In order to differentiate between the performances of these models, the data presented in **Figure 3.2** and **Table 3.2** is further analyzed by means of the absolute error (see **Table 3.3**) between the experimental data and the predictions of each distribution, with regard to Mean, Standard Deviation (STD), Kurtosis (K) and Skewness (S) values. It is apparent from **Table 3.3** that each distribution has its own strengths and weaknesses. The Normal distribution exhibits the largest discrepancy (absolute error) with regard to the mean, while the Poisson distribution performs very poorly with regard to standard deviation. With respect to Skewness, for both RAMLA and FBP, the absolute error is an order of magnitude smaller for the Gamma distribution compared to Poisson, Negative Binomial, Normal, Log-Normal distributions. With regards to Excess Kurtosis, for both RAMLA and FBP reconstructions, the discrepancy values are relatively small for all distributions except the Log-Normal distribution. For both RAMLA and FBP no model emerges as clearly superior, but Poisson is decidedly inferior.

Table 3.2: Cumulants and respective Margins of Error for 95% confidence level (in parentheses), calculated using maximum likelihood estimates for the parameters of Poisson, Negative Binomial, Normal, Log-Normal and Gamma distributions. Skewness and Excess Kurtosis of Normal distribution are always zero (follows from respective definitions)

Distribution	Mean	STD	Skewness	Kurtosis
RAMLA, 1st time frame of slice 21				
Poisson	2505	50.05	0.01998	0.0003992
	(2)	(0.02)	(0.00001)	(0.0000004)
Neg. Binomial	2505.0	203	0.34	0.13
	(0.2)	(7)	(0.02)	(0.02)
Normal	2498.5	204.2	0	0
	(0.2)	(0.1)		
Log-Normal	2505	204	0.245	3.107
	(11)	(8)	(0.009)	(0.008)
Gamma	2505	204	0.162	0.040
	(355)	(22)	(0.006)	(0.003)
FBP, 1st time frame of slice 21				
Poisson	3178	56.38	0.01774	0.0003146
	(3)	(0.02)	(0.00001)	(0.0000003)
Neg. Binomial	3178.3	310	0.55	0.32
	(0.5)	(11)	(0.04)	(0.04)
Normal	3174.1	311.6	0	0
	(0.3)	(0.2)		
Log-Normal	3178	312	0.30	3.16
	(17)	(13)	(0.01)	(0.01)
Gamma	3178	310	0.195	0.057
	(450)	(33)	(0.007)	(0.004)

Table 3.3: Comparison of discrepancy (absolute error) between experimental and calculated distributions (Poisson, Negative Binomial, Normal, Log-Normal and Gamma), with respect to Mean, Standard Deviation (STD), Skewness and Kurtosis (Excess Kurtosis) (see data from **Table 3.2** and **Figure 3.2**)

Distribution	Mean	STD	Skewness	Kurtosis
RAMLA, 1st time frame of slice 21				
Poisson	0	153	0.09	0.2
Neg. Binomial	0	0.2	0.2	0.03
Normal	6	0.9	0.1	0.2
Log-Normal	0.04	1	0.1	3
Gamma	0	0.3	0.05	0.1
FBP, 1st time frame of slice 21				
Poisson	0	254	0.2	0.05
Neg. Binomial	0	0.1	0.4	0.3
Normal	4	1	0.2	0.1
Log-Normal	0.07	2	0.1	3
Gamma	0	0.1	0.02	0.01

The poor representation of both RAMLA and FBP reconstructed PET data by the Poisson distribution is also shown using the Q-Q plots of **Figure 3.3.** Here, with the exception of minor deviations at the tails, the data points conform very well to the reference 45^0 lines for the Negative Binomial, Normal, Log-Normal and Gamma distributions. Significant differences in dispersion (and hence noise) are observed with regard to the Poisson distribution for both RAMLA and FBP.

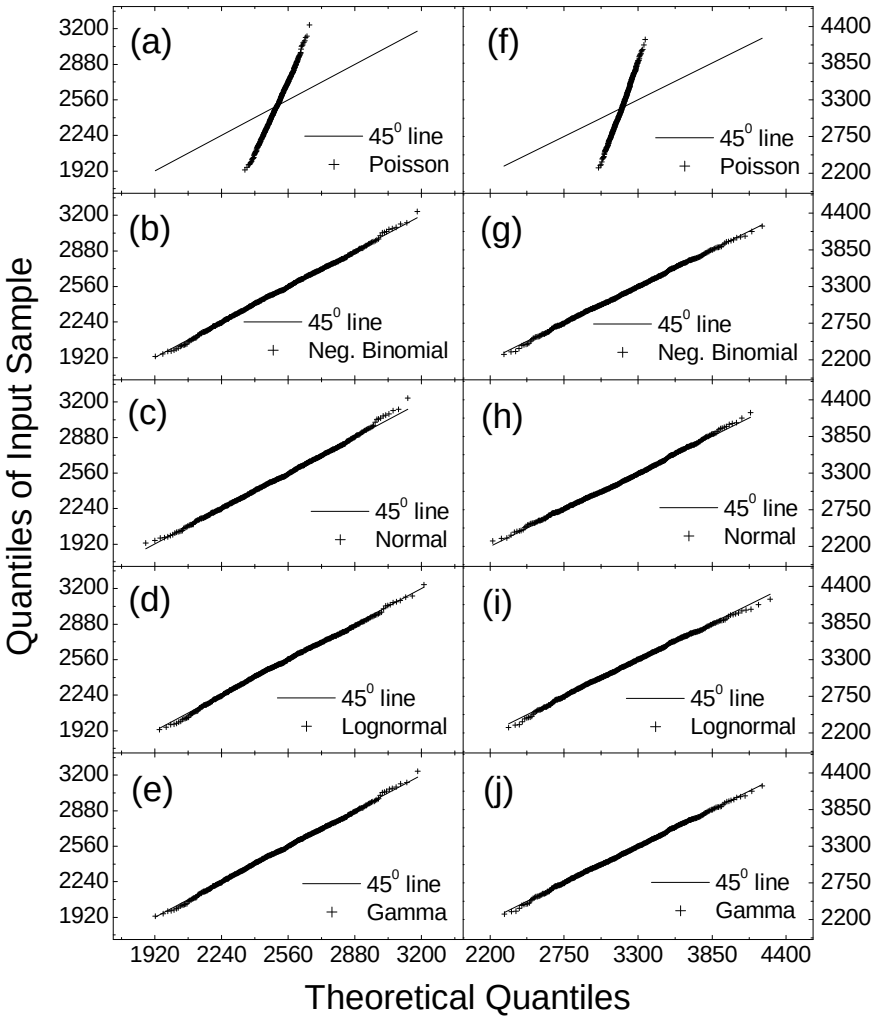

Figure 3.3: Q-Q plots (1st time frame of slice 21): circular ROI was selected on image reconstructed with RAMLA (a through e) and FBP (f through j)

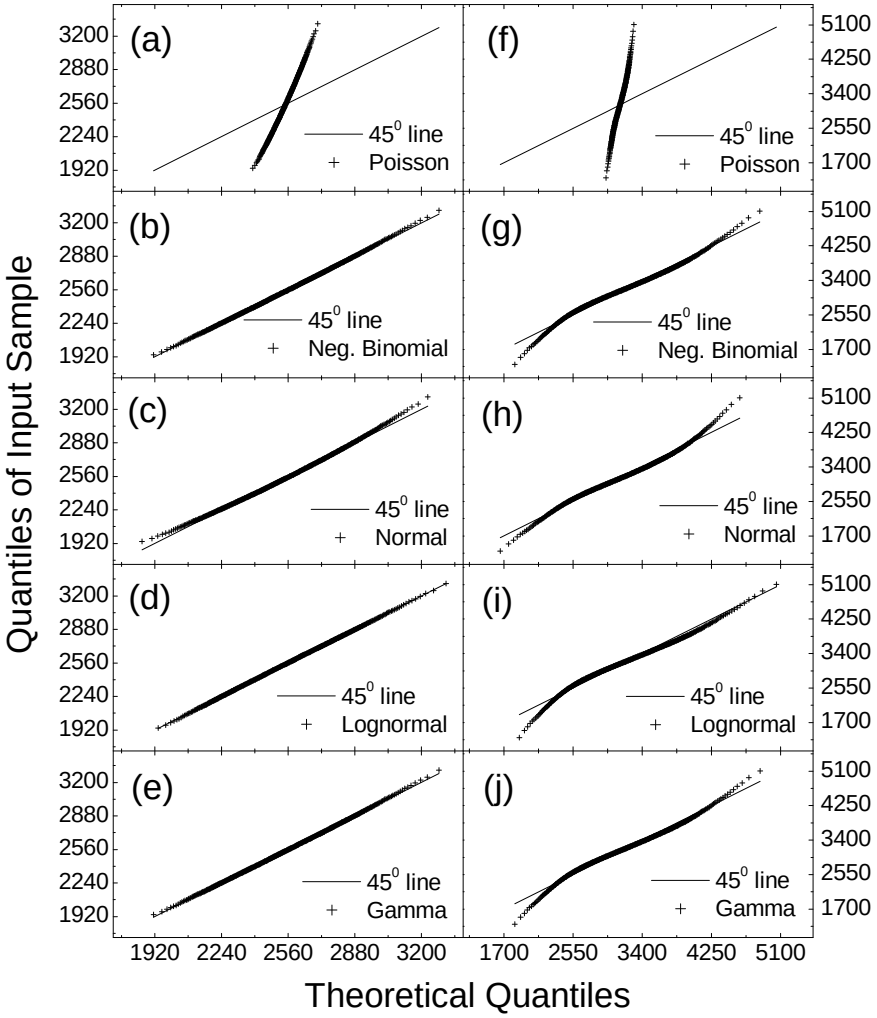

Figure 3.4: Q-Q plots: cylindrical VOI was selected on images of the first time frame for slices 4 to 42 reconstructed with RAMLA (a through e) and FBP (f through j)

The results presented thus far (**Table 3.3** and **Figure 3.1** through **Figure 3.3)** are representative only of the 1st time frame of slice 21. More general results may be sought by examining the data from all slice locations. Unfortunately, a full volumetric analysis is not possible due to the large changes in mean and standard deviation associated with slices nearest the minimum and maximum longitudinal extents (slices 1, 2, 3, 43, 44 and 45). The data from the full complement of slice

locations combine to produce an asymmetric quasi-bimodal distribution. Limiting the volumetric analysis somewhat arbitrarily to the combined results from slices 4 through 42 (first time frame), over which the values of mean and standard deviation are relatively constant, yields the Q-Q plots shown in **Figure 3.4** which reflect the conclusions drawn for the 2D case of slice 21 above.

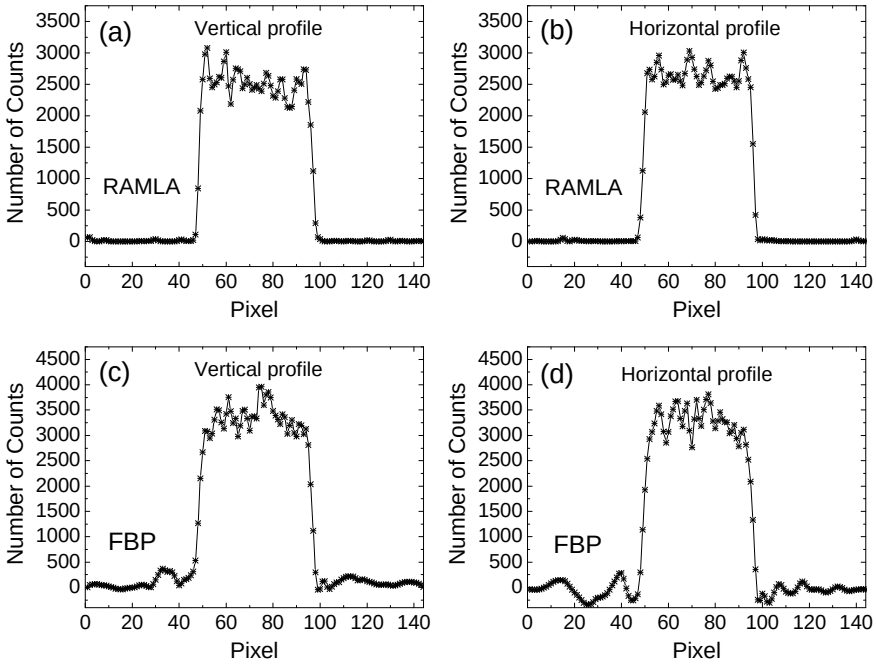

Figure 3.5: 3D PET study on cylindrical phantom (Slice # 21). 1D horizontal profile through the center of image reconstructed using RAMLA (upper right) and using FBP (lower right). 1D vertical profile through the center of image reconstructed using RAMLA (upper left) and using FBP (lower left)

Moreover, the Q-Q plots comparing experimental data to the Negative Binomial (**Figure 3.4b)**, Log-Normal (**Figure 3.4d)** and Gamma (**Figure 3.4e)** distributions reveal points that lie almost exactly along the 45^{o} reference line for RAMLA data. Deviations observed at the tails are somewhat more pronounced for the images

reconstructed with FBP (**Figure 3.4g through Figure 3.4j**). Poisson statistics provides, once again, an inferior description of the experimental data.

The spatial distribution of noise within the phantom was initially investigated along single pixel width vertical and horizontal diametric profiles. The magnitude and spatial distribution of noise in PET is significantly affected by the reconstruction algorithm used, as can be seen in **Figure 3.5** for the vertical and horizontal diametric profiles of Slice #21. Consider first the data provided by RAMLA reconstruction (**Figure 3.5a and Figure 3.5b)**. Noise appears to be independent of lateral position within the active volume. Outside the active volume, count values fall off rapidly to a nominal zero level. A degree of ambiguity attends this assessment, however, due to the significant noise evident in the data. A different picture emerges from the FBP data of **Figure 3.5c** and **Figure 3.5d**. As noted previously, count levels within the active volume are higher than their corresponding values provided by RAMLA reconstruction. Further, count values appear moderately peaked near the center of reconstruction, falling off slightly toward the outer regions of the active volume. The transition from active volume to cold outer regions is more gradual with FBP in comparison to RAMLA and regions of unphysical negative counts are observed external to the active volume.

Five circular sub-regions (top, bottom, left, right, and center) within the area of interest were analyzed. Each of these sub-regions had an area equal to ≈9 % of the transaxial cross section of the phantom. These sub-regions were propagated throughout slices 1 through 45 of the uniform cylindrical phantom resulting in the data of **Figure 3.6**. Immediately obvious are the longitudinal variations in total counts and noise within each region of interest. Rising rapidly from artificially low mean count values at the extrema slices (1 and 45), the RAMLA data of **Figure 3.6a** initially overshoots (slices 3 and 43) prior to assuming more nominal values characteristic of central slice locations.

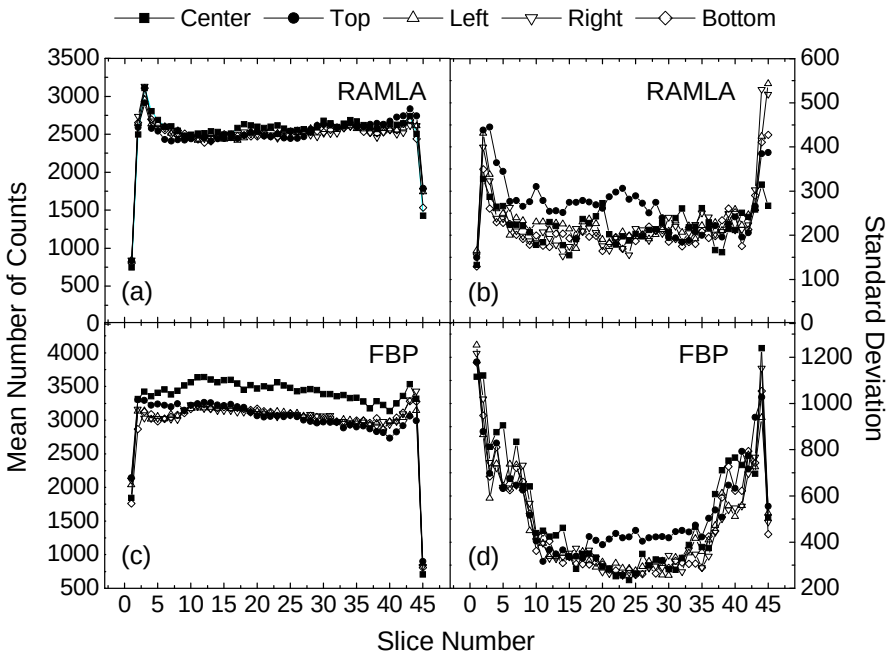

Figure 3.6: 3D PET study on cylindrical phantom (Slices 1 to 45); circular ROIs were selected on the phantom (Center, Top, Left, Right, and Bottom), mean number of counts and standard deviation was determined for each ROI

The unphysical diminishment of mean count values is greater at the superior extrema (slice 1) as compared to the superior most image location (slice 45). Further, a greater overshoot also occurs in the superior portion of the scan in contrast to its inferiorly located counterpart. From slice 9 through to 39 an overall upward trend in mean counts is observed in transiting from superior to inferior longitudinal positions. Noise levels (SD) with RAMLA reconstruction (**Figure 3.6b**) are lowest at the superior most position, spike rapidly at slice #2 and then quickly decreases over the next few slice locations. At the inferior most location noise levels all rise in comparison to central longitudinal locations. Mean count levels reconstructed with FBP (**Figure 3.6c**) are also lowest at the extrema locations. In contrast to the RAMLA data, the lowest FBP mean count level is found at the inferior most slice location. A modest inferiorly located overshoot is

again observed while the large superior overage seen with RAMLA reconstruction is completely absent with the FBP data. A gradual decrease in mean counts is observed as one progresses inferiorly from slice 12 to 39, which is opposite to the trend observed with RAMLA. At all but the last two slice locations, the greatest mean count level is maintained by the central sub-region. Noise levels with FBP reconstruction (**Figure 3.6d**) are lowest at mid-longitudinal locations ranging from slice 11 to 24 for the top and centre sub-regions respectively. Proceeding outward from these locations, noise levels climb to reach maximum values at slice #1 and #44. Noise levels drop precipitously in the transition from slice 44 to slice 45 at the inferior most extent of the reconstructed data. These variations observed over the longitudinal extent of the data for both reconstruction schemes result from the interplay of the inherent sensitivity profile of the scanner and the corrections algorithms which are applied. The proprietary nature of the finer points of the algorithms employed, and the manner in which they are implemented, renders moot further insight into the longitudinal structure of the data observed.

An evaluation of Normalized Standard Deviation (STD/mean), Skewness, and Excess Kurtosis over slices 4 through 42 (first time frame) is presented in **Figure 3.7** (RAMLA: a through c; FBP: d through f). With regard to normalized standard deviation (**Figure 3.7a** and **Figure 3.7d**) it may be seen that all but the Poisson distribution demonstrate good conformity to the experimental data. For FBP the Log-Normal distribution demonstrates a slight discrepancy at the extrema slices. When considering Skewness (**Figure 3.7b** and **Figure 3.7e**), the Negative Binomial distribution produces the least agreement, especially for the extrema slices. Analysis with respect to excess Kurtosis is presented in **Figure 3.7c** and **Figure 3.7f.** All distributions, except Log-Normal, model the RAMLA data reasonably well. The Log-Normal distribution also produces relatively poor agreement with FBP data except at the extrema slice locations where the negative binomial distribution yields even worse results.

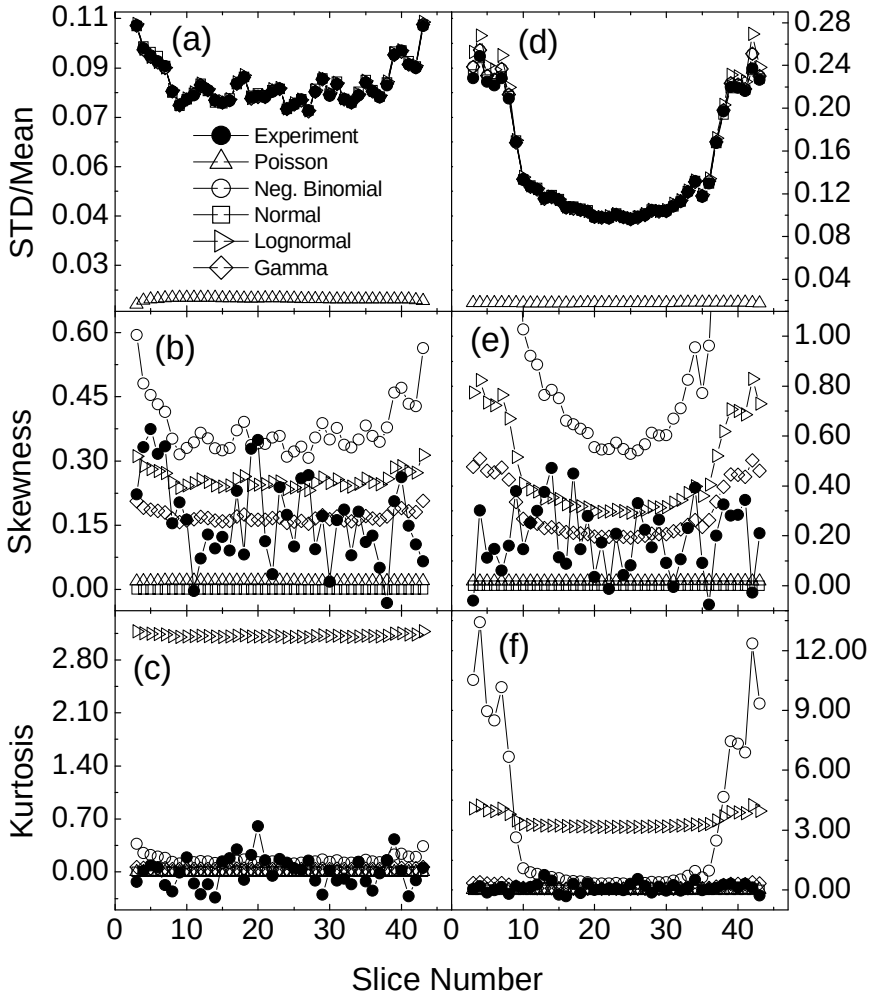

Figure 3.7: 2D case (1st time frame): STD/Mean, Kurtosis and Skewness with respect to experimental data. ROI was selected on images (slices 4 to 42) reconstructed with RAMLA (a through c) and FBP (d through f)

Figure 3.8 and Figure 3.9 show images from all time frames for slice 21 in the uniform phantom, reconstructed using RAMLA and FBP.

The dependence of statistical properties on count rate and counts collected is examined in **Figure 3.10** by investigating the time series of acquisitions. Here each data point represents the spatial integration over slices 4 through 42 in order to provide improved statistical accuracy. Temporal variability in the fit of the different models is clearly evident. This variability should not be surprising as all models are but approximations of reality built upon simplifying mathematical assumptions. Exact agreement between model and experiment is never guaranteed. **Figure 3.10a** and **Figure 3.10d** present the temporal development of normalized standard deviation (NSD).

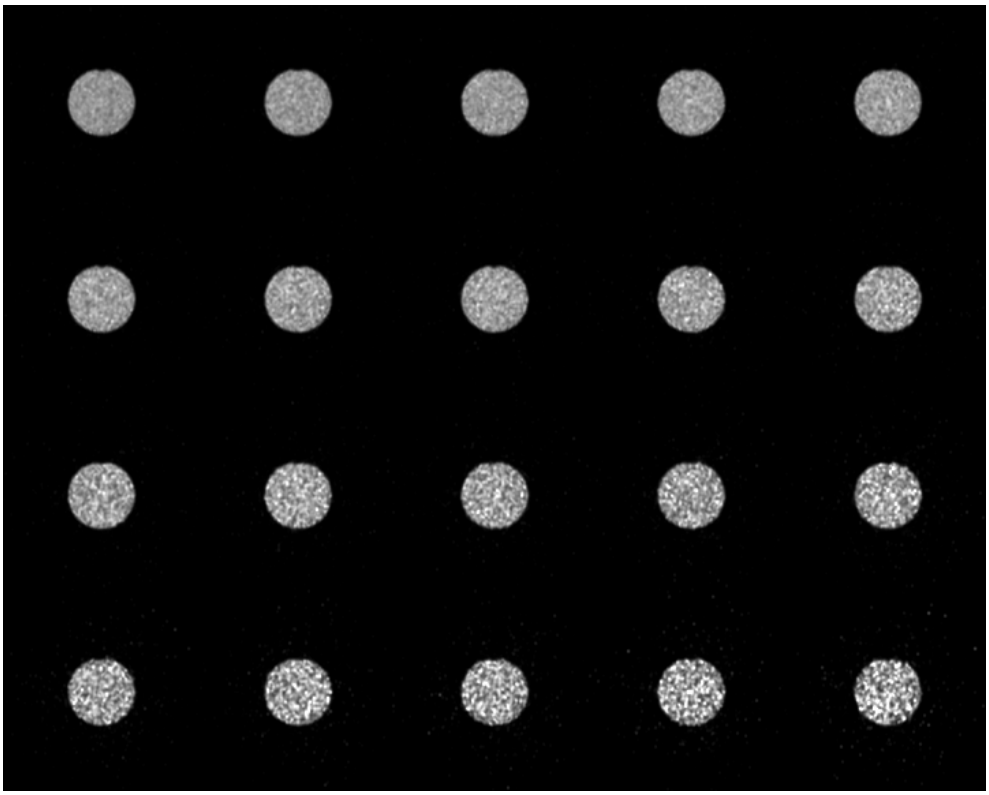

Figure 3.8: Slice 21 (all time frames, 1 to 20) reconstructed with RAMLA

As expected, the decreasing count levels associated with each subsequent time frame result in a monotonic increase in NSD (STD/Mean). With the exception of Poisson, good agreement is observed between all statistical models and the experimental RAMLA data for NSD. For this parameter, the Normal distribution provides good agreement over all time frames with the FBP data while all other statistical models diverge slightly from experiment as time increases. The Poisson distribution once again provides poorest overall agreement. Concerning Skewness, the RAMLA data shows an overall increase as the peak in its histogram shifts toward lower count values as a function of time.

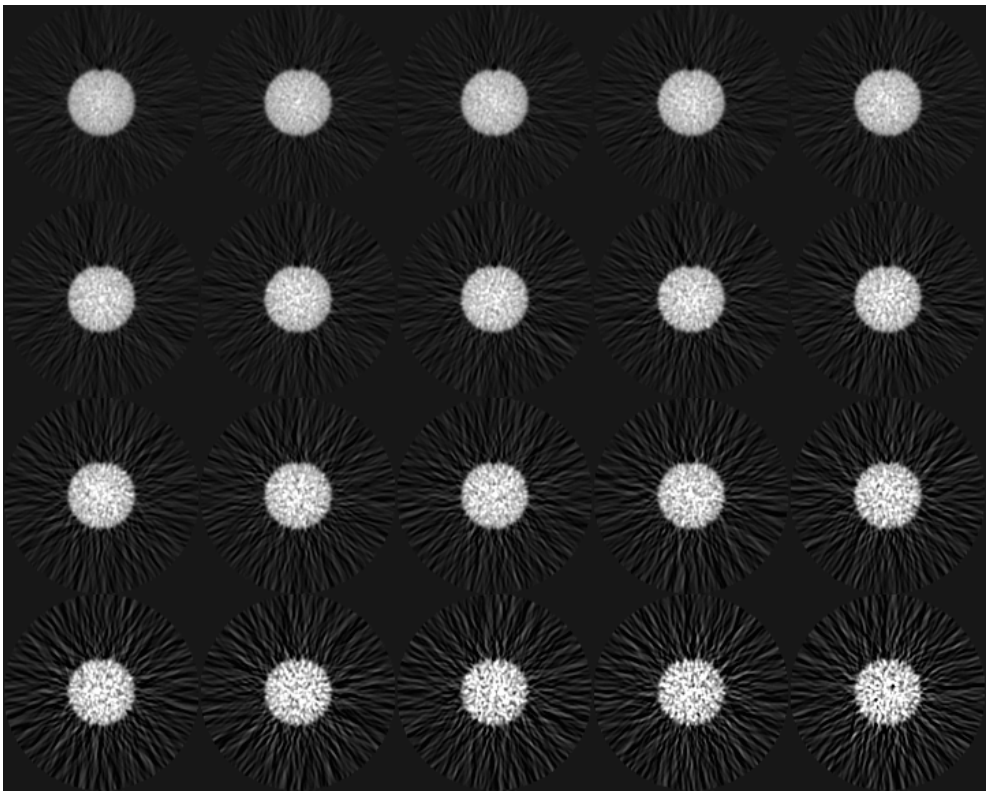

Figure 3.9: Slice 21 (all time frames, 1 to 20) reconstructed with FBP

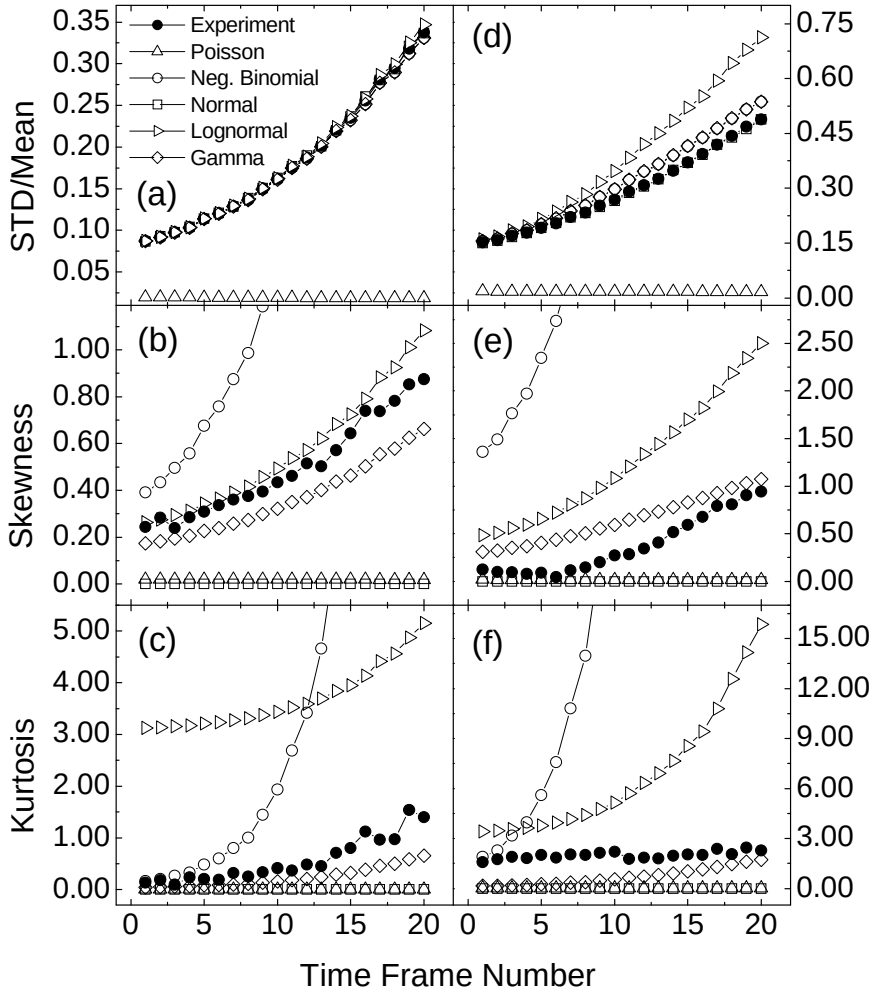

Figure 3.10: 3D case: STD/Mean, Kurtosis and Skewness with respect to experimental data. Cylindrical VOI was selected on images (slices 4 to 42) reconstructed with RAMLA (a through c) and FBP (d through f)

Skewness for both the Log-Normal and Gamma distributions follows the experimental data closely. FBP data shows an initial slight decrease in Skewness followed by a rise as time progresses further. Initially, the smallest discrepancy is seen with the Poisson and Normal distributions, while the experimental data conforms closest to the Gamma distribution for the latest time frames. FBP data

is modeled most poorly in regards to Skewness by the Negative Binomial distribution. With regard to excess Kurtosis, both RAMLA and FBP data demonstrate an overall trend of increase as time progresses. This increase in peakedness is a direct result of the shift in the histogram peak towards lower count values as indicated by the Skewness results. The ability of the Negative Binomial distribution to model experimental excess Kurtosis quickly diminishes after the first couple of frames. Overall, the Gamma distribution demonstrates the closest agreement with respect to excess Kurtosis for both RAMLA and FBP. The Poisson distribution proved least capable of modeling the experimental data as it completely fails to predict any temporal development of NSD, Skewness, or excess Kurtosis. The same is true for the Normal distribution with regard to Skewness and Excess Kurtosis. The Negative Binomial distribution demonstrates the most rapid divergence from experimental data as a function of time for the metrics of Skewness and excess Kurtosis. To provide a more quantitative description of the performance, with respect to the experimental data, for each of these models, the Root Mean Square Error (RMSE) is calculated for each metric in **Figure 3.10** and is presented in **Table 3.4**.

Table 3.4: 3D case (all time frames (1 to 20), VOI: slices 4 to 42), Root Mean Square Error (RMSE) for NSD, Kurtosis and Skewness with respect to experimental data is compared between different models for data reconstructed with RAMLA and FBP

	Metric	NSD		Skewness		Excess Kurtosis	
	Reconstruction	RAMLA	FBP	RAMLA	FBP	RAMLA	FBP
Model	Poisson	0.2	0.3	0.5	0.5	0.7	2
	Neg. Binomial	0.003	0.03	2	8	12	117
	Normal	0.002	0.003	0.5	0.5	0.7	2
	Log-Normal	0.003	0.1	0.9	1	3	6
	Gamma	0.003	0.03	0.1	0.3	0.4	1

This table clearly shows the advantages and disadvantages of each statistical approach to modeling the experimentally reconstructed data. For RAMLA, the Gamma distribution clearly yields the closest overall conformance to the actual time evolution of the experimental data, followed by the Normal distribution. It should also be noted that for temporal development of NSD, all distributions in question yield very similar values for RMSE when compared to experimental data. At the same time for Skewness and excess Kurtosis, the RMSE values for Negative Binomial and Log-Normal distributions are an order of magnitude higher than for Normal and Gamma distributions. For FBP, the Gamma distribution models very well Skewness and excess Kurtosis followed very closely by the Normal distribution while at the same time demonstrating greater discrepancy with regard to NSD than the Normal distribution. Both Negative Binomial and Log-Normal distributions fail to model the actual time evolution of the experimental data with respect to Skewness and excess Kurtosis, in comparison with Normal and Gamma distributions. NSD (STD/Mean) is modeled equally well by the Negative Binomial, Normal, Log-Normal and Gamma distributions for RAMLA and by the Negative Binomial, Normal, and Gamma distributions for FBP reconstructions. For FBP reconstruction, the Normal distribution demonstrates the least deviation from the time evolution of experimental data with respect to NSD while for the Log-Normal distribution the value of RMSE is very close to a Poisson distribution and two orders of magnitude higher compared to the Normal distribution.

3.3.3 Conclusions

In this chapter we have presented an investigation of the statistical properties of noise in PET images of a cylindrical phantom containing a uniform distribution of activity reconstructed with Filtered-Backprojection (FBP) and Row-Action Maximum Likelihood Algorithm (RAMLA), after all clinical correction and image reconstruction procedures have been applied. This analysis has shown that the noise in PET images created with RAMLA reconstruction is very well characterized by the Gamma distribution followed closely by the Normal

distribution, while FBP produces comparable conformity with both Normal and Gamma statistics. We have also shown that NSD (STD/Mean) is modeled equally well by Negative Binomial, Normal, Log-Normal and Gamma distributions for RAMLA and by Negative Binomial, Normal, and Gamma distributions for FBP reconstructions. While radioactive decay is well-modeled as a Poisson process, the net result after all correction and image reconstruction techniques have been applied is decidedly non-Poisson. This has important implications for an accurate evaluation of quantitative information provided by PET imaging. It is particularly true for dynamic PET imaging, where the signal-to-noise ratio decreases for each subsequent time frame, which can pose significant challenges for quantitative analysis. A large number of noise reduction techniques are predicated on additive noise models and an incorrect treatment of image noise can be detrimental to adequate algorithm performance. Noise reduction algorithms specifically designed for Poisson noise are expected to produce inferior results when applied to clinical PET images.

4. Noise Suppression Techniques in PET Imaging

PET images are, relatively speaking, excessively noisy in comparison to other imaging modalities (such as CT and MRI) which are commonly used in RT. Noise levels observed in PET images complicate their geometric interpretation. It is thus desirable to suppress PET image noise prior to segmentation. In our quest to denoise PET images, we have evaluated three techniques for their suitability as a preprocessing step prior to tumor volume delineation: (1) Principal Component Analysis (PCA)[49, 82], (2) Statistical and Heuristic Image Noise Extraction (SHINE)[67, 68], (3) Image Denoising Using Gaussian Scale Mixtures in Wavelet Domain (BLS-GSM)[83].

4.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a form of unsupervised learning. It is a linear projection method used to reduce the number of parameters in a problem and can be viewed as a rotation of existing axes to new positions in a space defined by the original variables. These new orthogonal axes represent directions of maximum variability. PCA has been successfully applied in medical imaging for image denoising and feature extraction[49, 82].

PCA is a data driven technique which is used to identify the variance-covariance structure in a data set. PCA utilizes the Karhunen-Loève theorem (theory of stochastic processes) which states that a stochastic process can be expressed as a linear combination of orthogonal functions. This is similar to Fourier series decomposition of a function. Whereas a Fourier series uses real numbers as coefficients for expansion into a basis of sine and cosine functions, the Karhunen-Loève transform employs random variables as coefficients for an expansion into a basis consisting of orthogonal functions which are determined by a covariance function of the process. The discrete Karhunen-Loève Transform (KLT), where coefficients are obtained empirically from the data sample, is known as Principal Component Analysis (also known as Proper Orthogonal Decomposition (POD), Hotelling Transform and Principal Component Transform). PCA, as applied to a dynamic PET data set, can be used for denoising

and identification of structures. Application of PCA to a multivariate image results in a new transformed multivariate image. The images contained within this new data set, the principal component (PC) images, represent the orthogonal directions of maximum variance for the original data. Mathematically, these directions are derived as the eigenvectors of the covariance matrix for the data set.

Each slice in a dynamic PET data set consists of *n* registered time frames, meaning that there are n pixels for every (*i, j*) coordinate point, one pixel at that point for each frame. These pixels may be arranged in the form of a column vector ***x*** (**Figure 4.1**). For an *M*×*N* image, there will be a total of *MN* such *n*-dimensional vectors comprising all pixels in the *n* time frames[84].

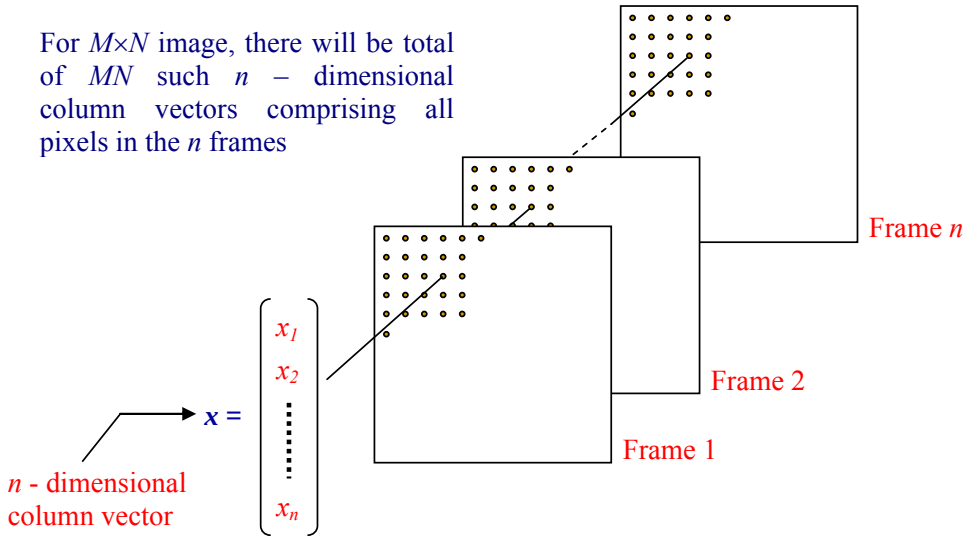

Figure 4.1: Forming a vector from corresponding pixels in a stack of time frames for a given voxel in a PET data set. (Adapted from *Gonzalez et al.*[84])

The multidimensional image data matrix with *M* rows, *N* columns, and *n* variables is unfolded, forming a two-dimensional matrix of size *n(MN)*[84]. The *n*×*n* covariance matrix $\boldsymbol{C}_x$ of the mean corrected data set is given by [84]:

$$C_x = \frac{1}{K-1}\sum_{k=1}^{K}\left(x_k - m_x\right)\left(x_k - m_x\right)^T \qquad \textbf{4.1}$$

where $K = MN$ and m_x is a mean of the vector population $\boldsymbol{x}$, which can be approximated by the sample average[84]:

$$m_x = \frac{1}{K}\sum_{k=1}^{K} x_k \qquad \textbf{4.2}$$

The principal component transform is given by[84]:

$$y = A\left(x - m_x\right) \qquad \textbf{4.3}$$

where the rows of transformation matrix A are the normalized eigenvectors of $\boldsymbol{C_x}$. The eigenvalues of $\boldsymbol{C_x}$ form the main diagonal elements $\boldsymbol{C_y}$. The main diagonal element in the i^{th} row of $\boldsymbol{C_y}$ is the variance of vector element $\boldsymbol{y_i}$. It can be shown that the covariance matrix $\boldsymbol{C_y}$ is diagonal, meaning the elements of vector $\boldsymbol{y}$ are uncorrelated. The original image can be recovered by performing the inverse transform[84]:

$$x = A^T y + m_x \qquad \textbf{4.4}$$

The denoising properties of PCA can be explored by reconstructing the image using only $\boldsymbol{q}$ eigenvectors, in which case $\boldsymbol{A}$ becomes a $q \times n$ matrix $\boldsymbol{A_q}$[84]:

$$\hat{x} = A_q^T y + m_x \qquad \textbf{4.5}$$

The name Principal Component Analysis (or transform) comes from the fact that only eigenvectors corresponding to the largest (principal) eigenvalues of the covariance matrix are used in reconstruction. Each eigenvalue corresponds to a Principal Component (PC) image (channel). In this work only the first principal component (PC1) image corresponding to the largest eigenvalue is considered. The data in the first PCA channel is far less noisy and significantly sharper than in the original image[49, 82, 84].

The PCA algorithm[84] used in this evaluation was implemented with MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA).

4.2 Statistical and Heuristic Image Noise Extraction (SHINE)

SHINE (Statistical and Heuristic Image Noise Extraction)[67, 68] is a special software algorithm, which can be used to extract important information from an image and remove noise by means of Correspondence Analysis. To distinguish information from noise it uses an original "stop condition" such that at the end of the process, an image with reduced noise level is produced. Use of Correspondence Analysis (CA) in image processing was first proposed by *Marano*[85].

This method begins by dividing the image into small blocks of size $m \times m$ (m is generally equal to 4 or 8) and transforming the image into a matrix $\boldsymbol{X}$. In this matrix, each row corresponds to elements of a block and each column corresponds to a pixel in the blocks. For an $h \times h$ image divided into $m \times m$ blocks, matrix $\boldsymbol{X}(p,n)$ has $p = (h \times h)/(m \times m)$ rows and $n = m \times m$ columns. CA is then applied to the matrix X, aiming to reduce noise contained in the image. The same transformation is utilized in the JPEG (Joint Photographic Experts Group) compression technique except it uses a Discrete Cosine Transform (DCT) instead of CA.

The SHINE algorithm has a very simple structure. First, a correspondence analysis (CA) is performed on the matrix X and then the following procedure is applied to each block i (where $i = 1 \ldots p$)[67]:

(1) Factors k ($k = 1 \ldots n$) are sorted in descending order according to the square of their cosine with block i. The cosine between block i and factor k is a measure of the quality of the representation of the block i along k axis (*Lebart et al.*[67, 86]).

(2) Block i is reconstructed step-by-step. Factors k are included according to their rank, starting from the most significant factor.

(3) An "original stop condition test" is performed at each step.

The caveat here is the "original stop condition test". This is what makes SHINE different from the technique proposed by *Marano*[85], where compression (or denoising) is achieved by reconstructing the image from matrix *X* with only the first *q* factors. Essentially the same idea is used in denoising with PCA. The ambiguity is in how to choose the number of factors (or PC), required to achieve optimal results. In our studies with PCA, only the first factor (PC1) was kept. There are other commonly accepted guidelines, however, such as the "Kaiser Criterion" and "Scree Test" that, in practice, seem to yield acceptable results. In SHINE, this problem is solved by using the classical variance comparison test[67, 86, 87]. Block reconstruction is stopped when residual variance is significantly lower than noise variance. The goal is to recover only the signal variance in the reconstructed block, without including the noise variance. The stop condition at step *k* is given by the following:

$$\sigma^2_{residual}(k) < \frac{\sigma^2_{noise} \cdot \chi^2(df)}{df} \qquad \textbf{4.6}$$

where σ^2_{noise} is the noise variance in a block at step *k*; $\sigma^2_{residual}$ is the residual variance at step *k*; *df* is equal to (*n*-*k*-1); and $\chi^2(df)$ is given by the χ^2 table for *df* degrees of freedom and a 5% alpha risk.

Two fundamental assumptions are made here[67]:

(1) The variance of the *n* values in a block (initial variance, $\sigma^2_{initial}$) is the sum of the variance of the signal in the block (signal variance, σ^2_{signal}) and that of the noise variance (σ^2_{noise}) in the block ($\sigma^2_{initial} = \sigma^2_{signal} + \sigma^2_{noise}$).

(2) The noise in the block at step *k* follows Poisson statistics and an estimate of the noise variance in a block can be obtained as the mean of the *n* values in the block.

Following the first assumption, if the variance of the reconstructed block at step k is $\sigma^2_{reconstructed}(k)$, than it can be written that

$$\sigma^2_{initial} = \sigma^2_{reconstructed}(k) + \sigma^2_{residual}(k) \qquad \textbf{4.7}$$

The reconstruction of a block is completed when the condition given by **4.6** is satisfied and reconstruction of the next block begins, until every block comprising the original image has been denoised.

The denoising procedure described above is repeated $m \times m$ times in a sliding fashion to minimize artifacts arising due to division of the image into small blocks. The original image is sampled into blocks $m \times m$ times, with offsets ranging from 0 to $(m - 1)$ pixels in both the width (x) and the height (y) directions and the resultant denoised image is the mean of the $m \times m$ adaptively filtered images obtained with the different offsets in x and y.

In lieu of the absence of free and easy access to the original implementation of SHINE by *Hannequin et al.*[67], the algorithm was implemented with MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA). A procedure proposed in the original work[67] was used for validation of our implementation of the SHINE algorithm. A numerical line phantom (64×64 pixels) containing six horizontal lines and six vertical lines with increasing width (ranging from 1 to 6 pixels) was generated. In the noise-free image the background pixel value was set to 10, while a line pixel value was set to 20. Following this, seven simulated noise-free data sets were computed by multiplying the original noise-free image with the following scaling factors: 1, 2, 4, 6, 8, 10, and 20. Noisy acquisitions were created by adding Poisson noise generated with MATLAB function "imnoise". Finally, SHINE and a median filter were applied to the noisy images to obtain noise corrected data sets. Noise suppression using these two methods was evaluated and compared by means of Mean Normalized Error (MNE), which is defined as[67]:

$$MNE = \frac{100}{h \times h} \sum_{j=1}^{h \times h} \frac{abs\left(val\left(j\right) - ref\left(j\right)\right)}{ref\left(j\right)} \qquad \textbf{4.8}$$

where $h \times h$ is the number of pixels in the image, $val(j)$ is the value of pixel j in noisy or processed images and $ref(j)$ is the value of the pixel j in the corresponding noise-free images.

Figure 4.2 gives the value of MNE as a function of count level.

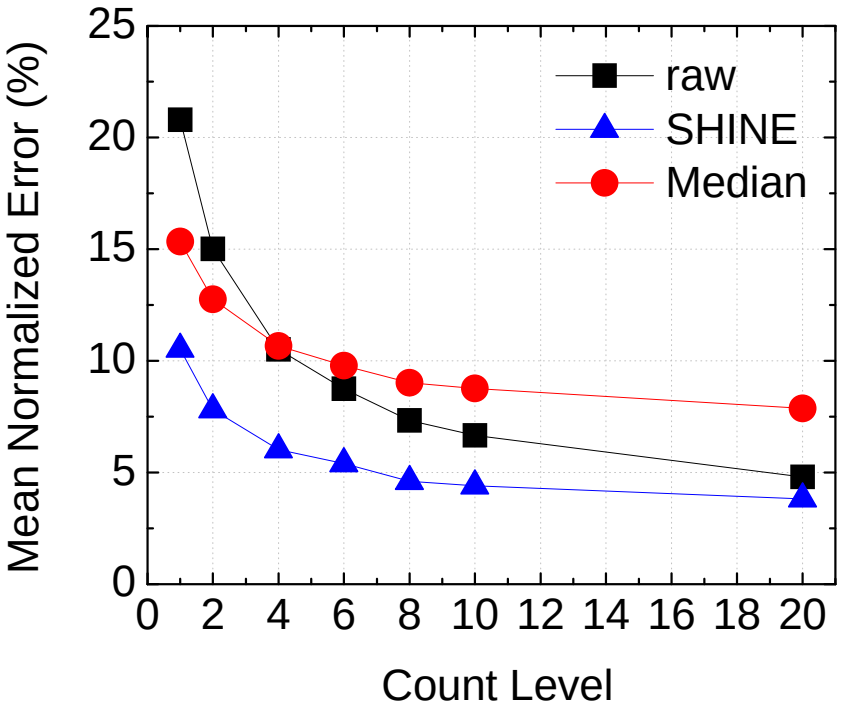

Figure 4.2: Quantitative results of noise extraction for a numerical line phantom. MNE as a function of seven count levels for raw images (squares), SHINE images (triangles) and median filtered images (circles). Reference images for the MNE calculations are the noise-free images.

Data presented in **Figure 4.2** closely replicates the results published in the original paper[67] and indicates that the MNE of SHINE images is small, compared

to the MNE of corresponding raw (noisy) images. In contrast to this the MNE curve for median filtering shows improvement only for low count (noise) levels. MNE curves for both SHINE and raw images follow an exponential form, suggesting that the relative noise reduction is constant and does not depend on the count level in the images. The mean ratio of raw MNE and SHINE MNE is 1.73, which is close to the value of 1.48 reported by *Hannequin et al.*[67]. The difference in values can be attributed to the statistical nature of the noise generation process.

In the original SHINE publication[67], the authors state that: "It can be applied to all scintigraphic images, *including PET data*, and to all low-count photon images" and "SHINE can be applied to all kinds of images: static images, *dynamic images*, whole-body images and SPECT or *PET images*". However, our work described in the previous chapter, concerning the statistical properties of noise in PET, shows that an assumption of strictly Poisson noise may lead to significant discrepancies. Other methods of noise estimation must be employed before SHINE can be applied to PET images. We will return to this point later in this chapter, after a discussion of BLS-GSM.

4.3 Gaussian Scale Mixtures in Wavelet Domain (BLS-GSM)

The choice of BLS-GSM (Bayes Least Squares - Gaussian Scale Mixture) was motivated by the fact that it has been shown to outperform some of the more recently published State-of-the-art methods[83, 88-92] as well as a number of standard well-known and widely-used denoising algorithms: a local variance-adaptive method in the pixel domain (implemented in MATLAB as function *wiener2*) and a hard thresholding method using an undecimated representation with five scales based on the minimum-phase Daubechies 8-tap wavelet filter[83]. BLS-GSM is an algorithm which uses scaled mixtures of Gaussian functions in the wavelet domain for image denoising[83]. It is based on a statistical model of the coefficients of an over-complete multiscale oriented basis. In this model, the collection of coefficients in neighboring positions is described as the product of two random variables, a Gaussian vector and a hidden positive scalar multiplier. The role of the multiplier is to modulate the local variance of the coefficients in the

neighborhood. This is done in order to account for empirically observed correlations between amplitudes of the coefficients. The BLS estimate of each coefficient is used to reduce the weighted average of the local linear estimate over all possible values of the hidden multiplier variable. This algorithm uses a procedure very popular in wavelet based denoising:

(1) Decompose the image into pyramid sub-bands at different scales and orientations

(2) Denoise each sub-band, except for the low pass residual band

(3) Apply an inverse pyramid transform to construct a denoised image

An important assumption made in BLS-GSM is that the image is corrupted by white *Gaussian noise* of *known variance*. In principle, BLS-GSM can also handle nonwhite noise of known covariance. In any case, before BLS-GSM can be used effectively in PET imaging, the algorithm has to be modified to allow for estimation of the standard deviation of the noise.

The original MATLAB implementation of BLS-GSM Denoising Toolbox, available on the internet from http://decsai.ugr.es/~javier/denoise (*Javier Portilla*), was used in this work.

4.4 SHINE(SIG) & BLS-GSM(SIG): A Modified Approach for Processing Gaussian Noise in PET Images

Both SHINE and BLS-GSM require specification of σ, the standard deviation of image noise. A hybrid approach was adopted for this purpose. Since σ is unknown a priori, it has to be estimated from the image. This is done by means of nonstandard decomposition of the two-dimensional discrete wavelet transform (DWT). Cascades of high and low pass filters are applied in order to decompose the image and produce a multi-resolution representation, where each wavelet coefficient represents the information content of the image at a certain resolution

in a certain position. In nonstandard decomposition, a wavelet transform is applied to an image by columns, then by rows, using the transform at one scale only. This technique produces a result in four quarters: the top left will be a half-sized version of the image and the other quarters are high-pass filtered images (see **Figure 4.3**). These quarters contain horizontal, vertical, and diagonal edges of the image. The next step is to apply a one-scale DWT to the top-left quarter, creating smaller images, and so on[84].

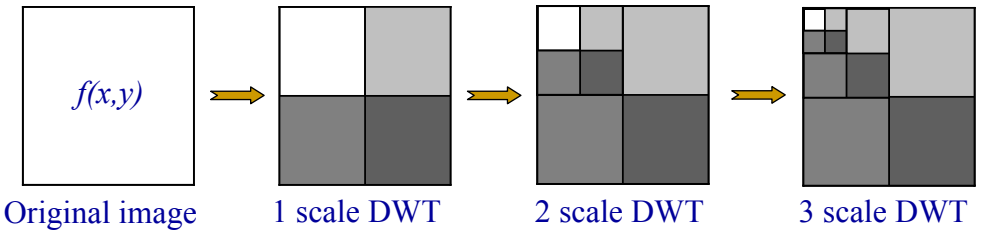

Figure 4.3: Nonstandard decomposition of the 2D DWT.

Donoho and Johnstone[93] proposed a robust noise level σ estimation based on the median absolute value of the wavelet coefficients as:

$$\sigma = \frac{median\left(abs\left(W\left[i\right]\right)\right)}{0.6745} \qquad \textbf{4.9}$$

where i is an index over all detail coefficients W at the finest decomposition level. The basic idea is that at the most detailed level, the wavelet coefficients are primarily due to noise. Due to the fact that only a few coefficients correspond to signal, use of the median acts to eliminate bias that might be otherwise introduced. The constant *0.6745* comes into play because for a set of independent Gaussian random variables $\{n_i\}$, i $\in \{1, 2 \ldots N\}$ with zero-mean and variance σ^2, E[median(abs(n_i))] $\approx 0.6745\sigma$. This estimator has become very popular in practice and is widely used[28, 54]. In estimating the noise variance we employed the Harr wavelet. The Harr wavelet is the simplest possible wavelet and it is the same as a Daubechies 2 tap wavelet. The number of taps corresponds to the number of filter coefficients. The Harr transform was in existence before Ingrid Daubechies introduced her class of wavelets, thus it retains its name. The use of a Daubechies

4 tap wavelet will produce smoother results than using the Harr wavelet. The Harr wavelet is not the best choice, since biomedical signals are characterized by smooth transitions as opposed to harsh jump discontinuities; however it allows the fastest computation of noise variance.

4.5 SHINE(SIG) & BLS-GSM(SIG): Comparison to State-of-the-Art Methods

Performance of the modified algorithms, SHINE(SIG) and BLS-GSM(SIG), was tested on a set of 8-bit grayscale test images, of size 512×512 and 256×256, each contaminated with additive Gaussian white noise with 10 different degrees of variance (noise levels 1 to 10 correspond to standard deviations of 1, 2, 5, 10, 15, 20, 25, 50, 75 and 100 respectfully) generated with the MATLAB function "imnoise".

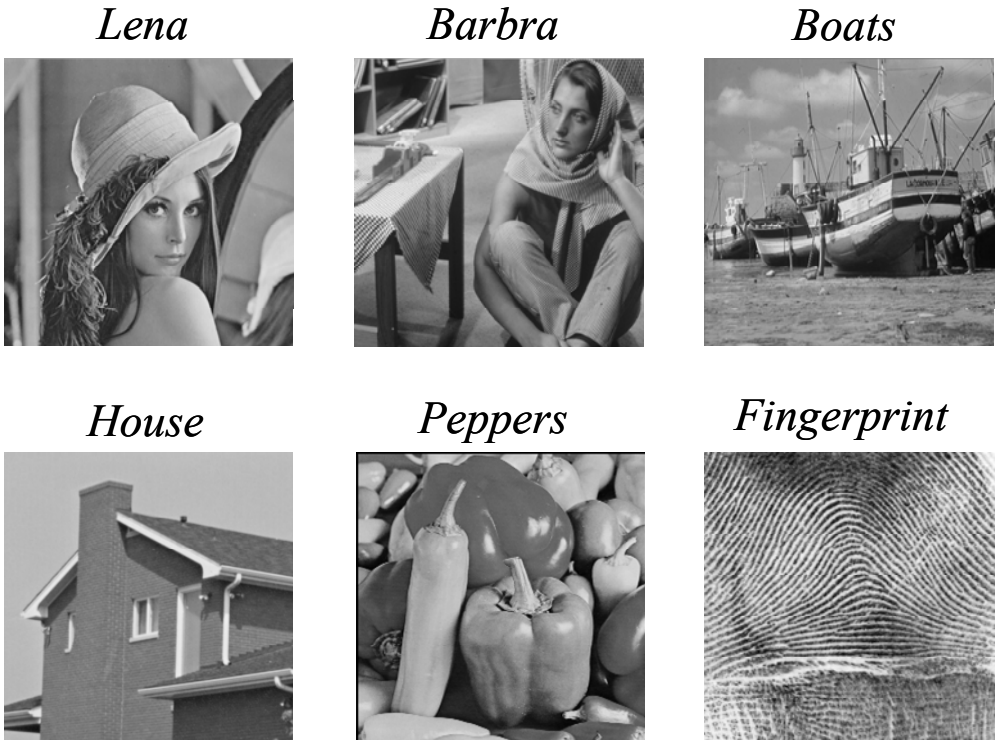

Figure 4.4: 8-bit grayscale test images used in SHINE(SIG) and BLS-GSM(SIG) algorithm performance evaluation

The commonly known Lena, Barbara, Boats, House, Peppers and Fingerprint images which are widely used in image processing literature (see **Figure 4.4**) were used here. To allow proper comparison between the algorithms, the same version of the images was used as in the original BLS-GSM paper[83]. A MATLAB implementation of the BLS-GSM Denoising Toolbox and 8-bit grayscale test images (**Figure 4.4**) are available on the internet from http://decsai.ugr.es/~javier/denoise (*Javier Portilla*).

For both SHINE and BLS-GSM noise-suppression algorithms, a required input is the standard deviation (σ) of noise in the image. Therefore, in our modified approach a Donoho and Johnstone[93] noise variance estimator is used (equation **4.9**). Since this estimator utilizes a wavelet transform, its performance may be potentially influenced by the choice of wavelet function. For this reason the performance of the estimator was evaluated for Harr and Daubechies 4 wavelets at 10 different degrees of variance (noise levels).

In **Figure 4.5** the average standard deviation (averaged across six test images generated from the images in **Figure 4.4**) of noise is plotted as function of noise level. It can be seen from **Figure 4.5**, that the performance of the estimator produces little difference between the Harr and Daubechies 4 wavelets. Therefore, it is reasonable to use the Harr wavelet for estimating noise variance. The Harr wavelet provides a fast estimation of noise variance and is used in all evaluations of SHINE(SIG) and BLS-GSM(SIG) algorithm performance described in this section.

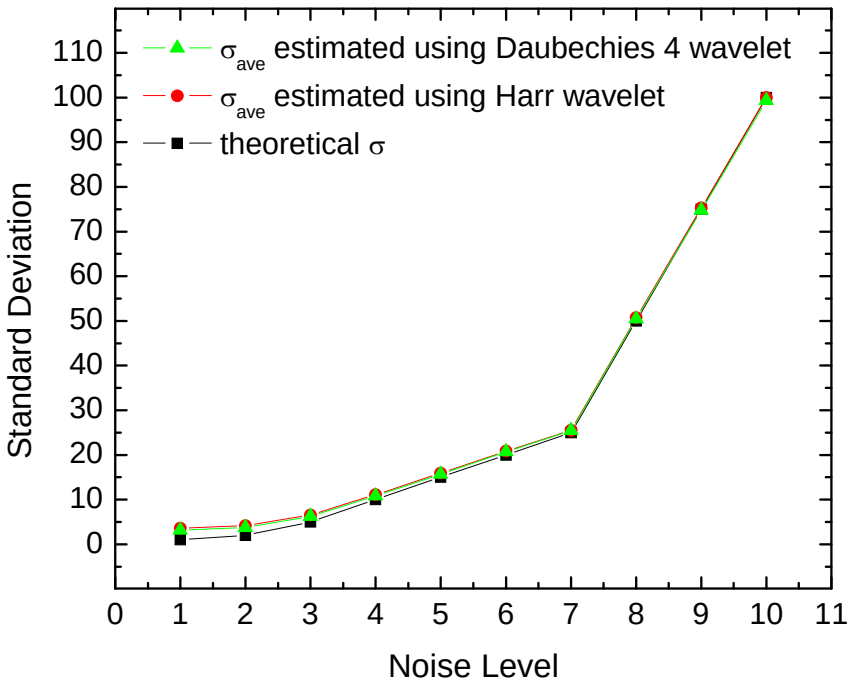

Figure 4.5: Performance evaluation of the Donoho & Johnstone estimator at 10 different contamination levels. Average standard deviation (averaged across six test images generated from images in **Figure 4.4**) of noise estimated at corresponding noise level is plotted as function of noise level. Noise Levels 1 to 10 correspond to standard deviations of 1, 2, 5, 10, 15, 20, 25, 50, 75 and 100 respectfully. Theoretical σ – standard deviation of Gaussian noise added to a test image at given noise level.

The relative performance of SHINE(SIG) and BLS-GSM(SIG) in denoising images contaminated with additive Gaussian white noise, with 10 different degrees of variance, was evaluated using several metrics: Pearson's Product-Moment Correlation Coefficient (PMCC), Peak Signal-to-Noise Ratio (PSNR), and Root Mean Square Error (RMSE).

Pearson's Product-Moment Correlation Coefficient (PMCC) between A and B is given by:

$$\mathrm{PMCC} = \frac{\sum_m \sum_n \left(A_{mn} - \overline{A}\right)\left(B_{mn} - \overline{B}\right)}{\sqrt{\left(\sum_m \sum_n \left(A_{mn} - \overline{A}\right)^2\right)\left(\sum_m \sum_n \left(B_{mn} - \overline{B}\right)^2\right)}} \qquad \textbf{4.10}$$

where $\overline{A}$=mean(A) and $\overline{B}$=mean(B) and A and B are $m \times n$ size monochrome images. PMCC estimates the strength of the linear relationship between two random variables and can take values between -1 (decreasing linear relationship) and 1 (increasing linear relationship). The closer the coefficient is to either −1 or 1, the stronger the correlation between the variables. To evaluate the performance of these denoising algorithms, two different values of Pearson's correlation coefficient are calculated at 10 different contamination levels. The first value, PMCC(IN), represents the correlation between original noisy and original noise-free images. The second value, PMCC(OUT), represents the correlation between denoised and original noise-free images. PMCC(OUT) as a function of PMCC(IN) is presented in **Figure 4.6** for the original BLS-GSM, BLS-GSM (SIG), and SHINE(SIG) with two different sampling window sizes.

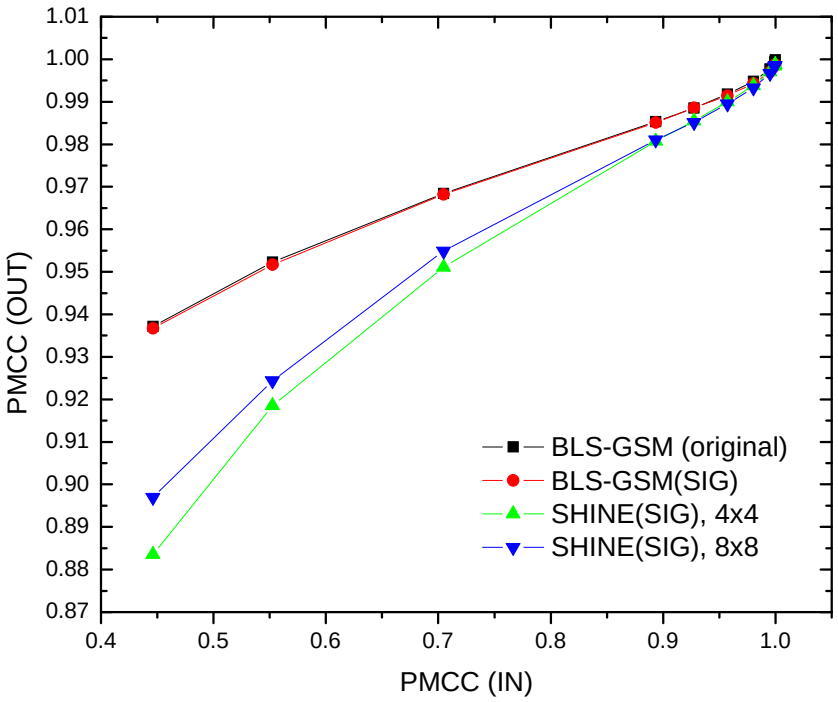

Figure 4.6: Performance evaluation of SHINE(SIG) and BLS-GSM(SIG) algorithms at 10 different contamination levels by means of Pearson's Product-Moment Correlation Coefficient (PMCC).

In **Figure 4.6**, in the case of SHINE(SIG), 4×4 and 8×8 refer to the size of the sampling window used. For the original BLS-GSM algorithm, the input variance of the noise was taken to be exactly equal to the variance of the generated white Gaussian noise, while for BLS-GSM(SIG) and SHINE(SIG) it was estimated using a Donoho and Johnstone[93] noise variance estimator (equation **4.9**). It is obvious from this figure that the introduction of the variance estimator into the algorithm did not change the performance of BLS-GSM (BLS-GSM(SIG) vs. BLS-GSM).

While the performance of both algorithms (BLS-GSM and SHINE) drops as noise levels increase, this decrease in performance is more dramatic for SHINE. It also appears that SHINE with an 8×8 window outperforms SHINE with a 4×4 window at very high noise levels. The PMCC values in this figure are averaged

across the six tested images. While Pearson correlation indicates the strength of a linear relationship between two variables, its value alone may prove insufficient for evaluation of this relationship. This is especially true in the case where the compared random variables are not normally distributed. Another drawback of the Pearson correlation coefficient is sensitivity to outliers.

To verify the results obtained using PMCC (**Figure 4.6**) two other popular metrics which appear in the signal processing literature may be employed. These are Peak Signal-to-Noise Ratio (PSNR) and Root Mean Square Error (RMSE). Both metrics can be defined via the mean squared error (MSE) given by:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[A(i,j) - B(i,j)\right]^2 \qquad \textbf{4.11}$$

where A and B are $m \times n$ size monochrome images corresponding to noisy (or denoised) and noise free images respectively.

The root mean square error (RMSE) is often used to quantify the differences model (or estimator) prediction and experimental observation.

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[A(i,j) - B(i,j)\right]^2} \qquad \textbf{4.12}$$

In this evaluation, the RMSE calculated between noisy and noise-free images is compared to RMSE calculated between images denoised with SHINE or BLS-GSM and the noise-free image (**Figure 4.7**). In **Figure 4.7** the Root Mean Square Error (RMSE) averaged across these six images is plotted versus the standard deviation of noise. In the case of BLS-GSM, the input variance of the noise was taken to be exactly equal to the variance of the generated Gaussian noise. It is clear from **Figure 4.7** that the introduction of a variance estimator into the algorithm did not change the performance of BLS-GSM (BLS-GSM(SIG) vs. BLS-GSM) and in the case of SHINE significant improvement is observed (SHINE (SIG) vs. SHINE). These results mirror conclusions drawn from data presented in **Figure 4.6**.

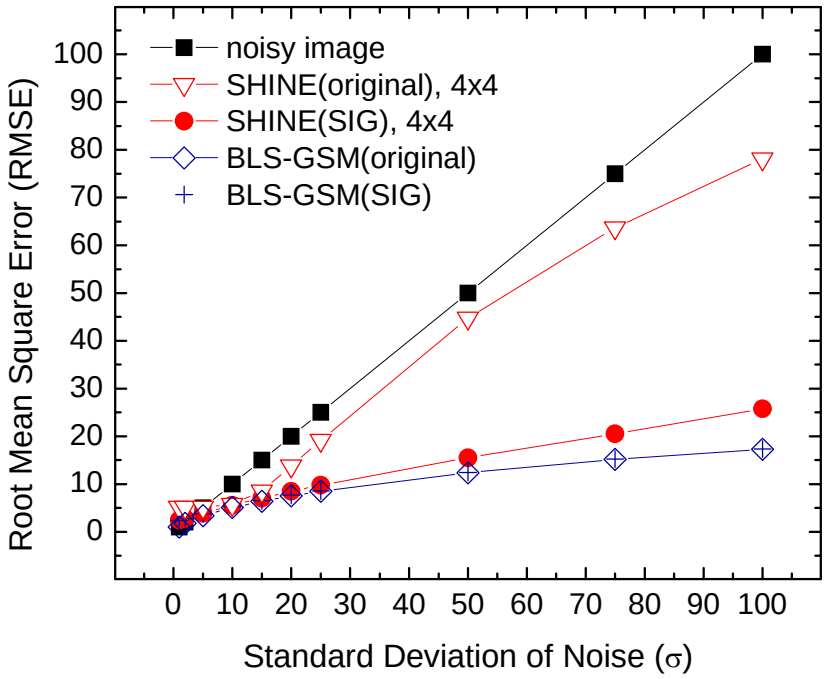

Figure 4.7: Performance evaluation of the SHINE(SIG) and BLS-GSM(SIG) algorithms at 10 different contamination levels by means of Root Mean Square Error (RMSE)

Finally, we have compared the performance of BLS-GSM(SIG) and SHINE(SIG) algorithms to the results of some more recently published state-of-the-art methods and these are shown in **Figure 4.8**. These methods are compared by means of peak signal-to-noise ratio (PSNR), calculated at 10 different contamination levels. The PSNR is widely used in engineering applications and represents the ratio between the maximum possible signal power and the power of corrupting noise that affects the fidelity of its representation. This is usually expressed in terms of the logarithmic decibel scale to accommodate description of signals with wide dynamic range.

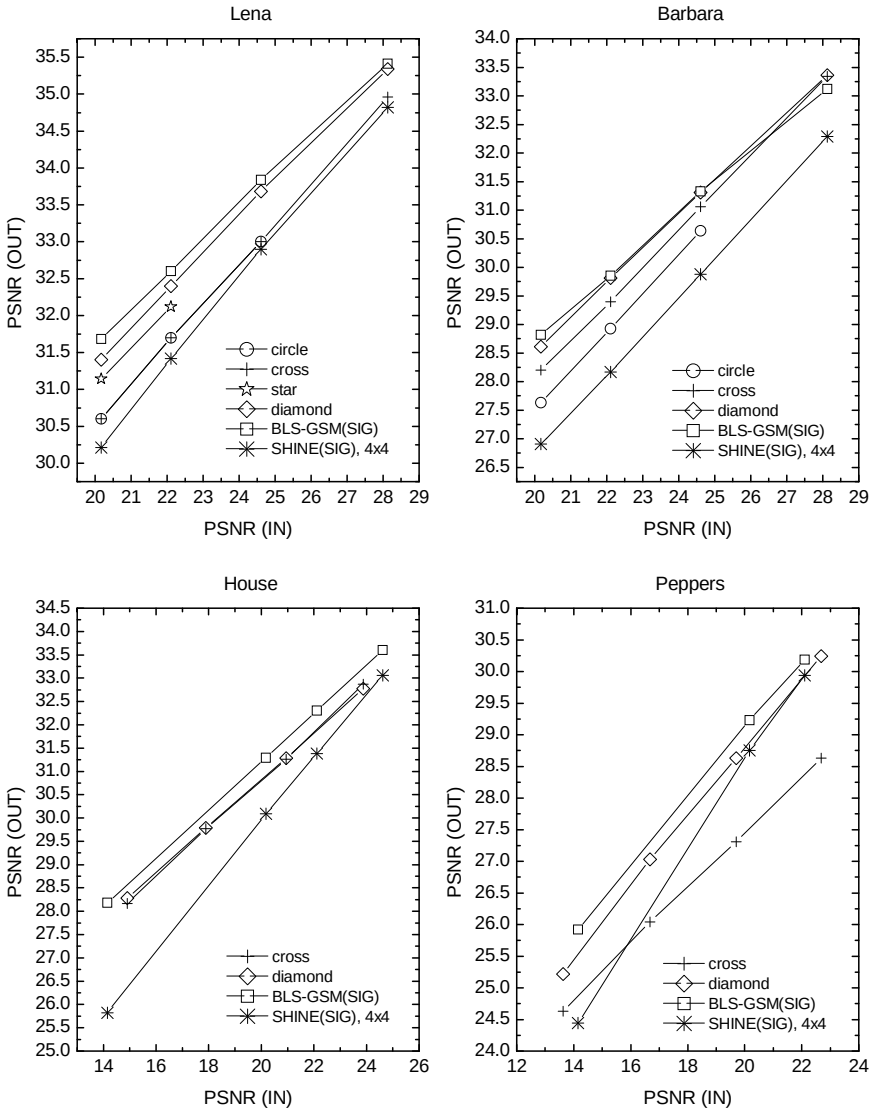

Figure 4.8: Comparison of the denoising performance of several best available published methods at four different noise levels with standard deviations of 10, 15, 20, and 25 (PSNR(IN) = 28.13, 24.61, 22.11, and 20.17 dB respectively). Square and asterisk symbols indicate BLS-GSM(SIG) and SHINE(SIG) results respectively. (Lena, Barbara) circles[89]; crosses[90]; stars[92]; (House, Peppers) crosses[88]; diamonds[91].

Peak signal-to-noise ratio (PSNR) can be calculated using the following expression:

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20 \cdot \log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \qquad \textbf{4.13}$$

where $MAX_I = 2^b$-1, and b is number of bits per sample. For instance if pixels are represented by 8 bits per sample, then $MAX_I = 255$.

In **Figure 4.8**, PSNR(IN), calculated between noisy and noise-free images, is compared to PSNR(OUT), calculated between denoised and noise-free images. In this figure, for the PSNR(OUT) resulting from denoising with the State-of-the-art methods, we have used data from previous work by *Portilla et al.*[83]. It is obvious from **Figure 4.8** that BLS-GSM(SIG) outperforms SHINE(SIG) and all the other methods compared in this evaluation.

4.6 Performance Evaluation: SHINE, SHINE(SIG), BLS-GSM(SIG), and PCA

In this analysis, the performance of SHINE, SHINE(SIG), and BLS-GSM(SIG) are compared to denoising with PCA by means of clinically relevant experimental phantom data. To facilitate this comparison, measurements were performed with a cylindrical phantom (long axis coincident with the reconstruction centre and orthogonal to the image plane; inside diameter 19.0 cm), which was filled with ^{11}C ($T_{1/2}$ = 20 min, A = 0.006 MBq/ml) to simulate background. A cylindrical target (inside diameter = 4.75 cm, inside length = 8.5 cm), filled with ^{18}F ($T_{1/2}$ = 110 min, A = 0.002 MBq/ml), was mounted in the centre of the phantom with its long axis coincident with the reconstruction centre to simulate a tumor volume. A dynamic sequence of 26 frames (45 slices per frame) was acquired for 130 min according to a schedule of 26 time frames of 300 s duration each. The data was corrected for attenuation and scatter, and then reconstructed using 3D-RAMLA (row-action maximum likelihood algorithm).

Figure 4.9 provides a visual comparison of a representative slice (Slice 33, first time frame) denoised using the original SHINE, SHINE(SIG), BLS-GSM(SIG), and PCA algorithms. Note that there is very little visual difference

between the noisy image and images denoised with SHINE(SIG) and BLS-GSM(SIG). The image denoised with the original SHINE algorithm appears slightly blurrier than with SHINE(SIG). This is due to the fact that in the original SHINE algorithm the noise variance changes between blocks (see section **4.2**) and is estimated as the mean of the *n* values in the block, while in SHINE(SIG) the noise variance is assumed to be constant throughout the image and is estimated using the Donoho and Johnstone[93] method (equation **4.9**).

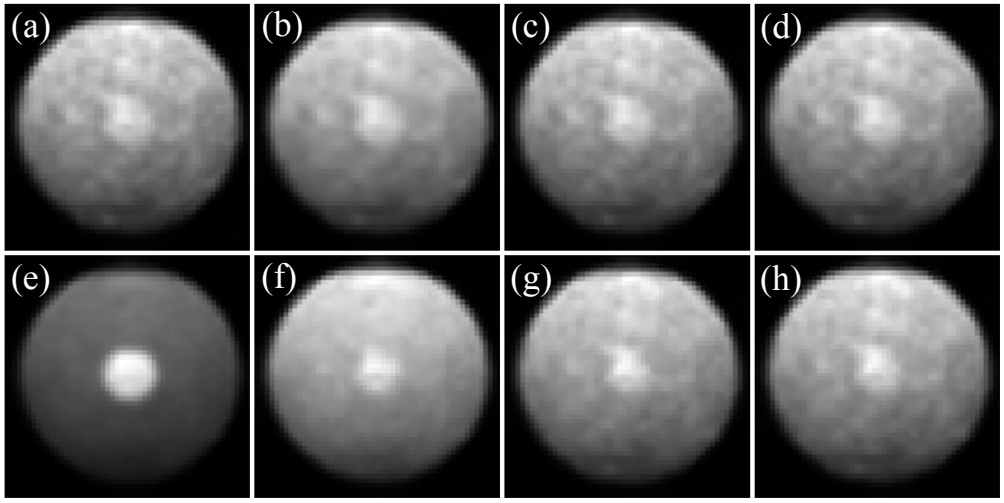

Figure 4.9: (a): Original Noisy Image (Slice 33, first time frame). Comparison of denoising performance of (b) SHINE; (c) SHINE(SIG); (d) BLS-GSM(SIG): and (e)-(h) PCA denoised images reconstructed using 1, 2, 3, and 4 eigenvectors.

Figure 4.9 (e, f, g, and h) depicts PCA denoised images reconstructed using 1, 2, 3, and 4 eigenvectors. It is evident from this figure that noise levels increase as more eigenvectors are used in PCA reconstruction and that PCA demonstrates superior performance in comparison with SHINE and BLS-GS. The greatest level of noise suppression is achieved in images reconstructed with only the first Principal Component (PC1, **Figure 4.9e**). Another advantage of using PCA is that, unlike SHINE and BLS-GSM, there is no need to process each image of a 4D PET dataset separately; making the overall processing time for PCA based

noise suppression significantly shorter as compared to SHINE and BLS-GSM. In this particular example, for PCA denoising, it took only 68.2 seconds for the entire data set (45 slices × 26 frames = 1170 images), while for SHINE, SHINE(SIG) and BLS-GSM(SIG) the processing times were 16.6, 11.7 and 9.7 seconds respectively for the Slice 33 of the first time frame alone. These lengthy processing times make SHINE, SHINE(SIG) and BLS-GSM(SIG) absolutely undesirable in applications where denoising of dynamic PET data is required due to the large number of images involved. Also, as demonstrated in **Figure 4.9e**, retaining only the first principal component (PC1) allows one to obtain a data set with a much lower level of noise while enhancing contrast, thereby optimizing the signal-to-noise ratio (SNR).

4.7 Noise Suppression Techniques: Conclusions

In this chapter, a comparison between the denoising of PET images with different techniques (PCA, BLS-GSM, and SHINE) was presented. This analysis suggests that PCA is a suitable approach for post-processing of dynamic PET data aimed at noise reduction prior to tumor volume delineation. PCA applied to dynamic PET image data is used to generate images with higher quality and improved SNR without modeling assumptions. Within the framework of the proposed algorithm, retaining only the first principal component (PC1) allows one to obtain a data set with a much lower level of noise while enhancing contrast, thereby optimizing signal-to-noise ratio (SNR).

5. Segmentation Techniques in Medical Imaging

Recent advances in Nuclear Medicine imaging modalities in general and in multimodality imaging (PET/CT & PET/MR) in particular have resulted in significant improvements in the areas of anatomical, functional, and dynamic imaging. With these advancements PET/CT and PET/MR based radiation therapy treatment planning have become an attractive reality. A crucial step in complementing CT or MR data with PET is the accurate delineation of cancerous tissues (or target volumes) on PET images and separating these from normal tissue for further processing. This procedure is known as image segmentation.

5.1 Image Segmentation

Image segmentation is an analysis technique used in image processing to subdivide an image into the components or objects comprising it, based on one or more characteristics. Most segmentation algorithms can be classified as a member of one of two categories: (1) methods based on an evaluation of *differences* in intensity values, (2) methods based on an evaluation of *similarities* in intensity values[94, 95]. In the first case, an image is segmented based on distinct changes in intensity, e.g. edges[94-96]. Popular choices are the Sobel edge detection technique, the Prewitt gradient based method and the Canny edge detector[94, 96]. Thresholding, clustering, and region growing are examples of methods falling into the second category. These are based on segmenting an image into regions of similarity according to a set of predefined criteria[94, 95]. It has been shown that improvement in segmentation is possible, if a combination of methods from the two categories or the same category is used[84, 95]. Examples are techniques in which edge detection or region growing is combined with thresholding[84, 95].

An automated segmentation algorithm is stopped when a desired level of detail is achieved, i.e. all the objects or regions of interest have been detected. This is one of the most difficult and non trivial tasks in modern image processing. Segmentation accuracy determines the success or failure of computerized analysis[95]. Many approaches have been suggested for image segmentation, but only a handful of those can be successfully applied to PET imaging[94, 95, 97].

5.2 Current Practices for Tumor Volume Delineation

In Nuclear Medicine Imaging, many segmentation approaches use the gray scale pixel values derived from a single image or image volume. A few of these methods, along with their advantages and shortcomings, are briefly described below.

Scans in PET can be acquired in two modes, static and dynamic. Static scans account for the majority of clinical procedures performed. In this mode the activity of the radiotracer is counted over a single fixed time period, resulting in a single image per anatomical slice location. At the present time these scans are analyzed visually by a radiation oncologist for the purpose of tumor volume delineation[32, 98-110]. The visual interpretation method is susceptible to image display Window/Level settings and has been shown to lead to significant inter-observer variability when comparing contours drawn by different radiation oncologist on the same PET/CT data[18, 32, 109, 111, 112]. As a result, the search for more objective approaches has begun [18, 32, 109, 111, 112]. An increasing number of emerging automated tumor volume delineation algorithms is being developed for static PET based imaging. These are mostly based on either a standardized uptake value (SUV) of 2.5 around the tumor[101, 110, 113, 114], a fixed threshold of maximum signal intensity (for example 40% of SUV_{max})[16, 19, 110, 115-119], or a variable threshold based on the maximum SUV which is dependent on the signal-to background-ratio (SBR)[32, 101, 110, 120-122]. A recent review published by John A. Lee[33] summarizes many delineation techniques popular in the literature.

Image segmentation using histogram based thresholding is by far the simplest and most popular technique, since it is easy to implement and requires less CPU (Central Processing Unit) time to run. The goal in this approach is to find the optimal threshold (or thresholds) allowing accurate classification of regions in an image. This is achieved through minimization (or maximization) of a criterion function based on the histogram of the image. The optimal threshold corresponds to the gray level value at which the criterion function attains its minimum (or maximum) value. Images can be segmented by choosing a threshold based either on an intensity level histogram of the entire image or using local information by

dividing the image into several sub-regions and calculating thresholds for each sub-region[97]. In the first instance, the process is called global thresholding. If several sub-regions are isolated in the image and a threshold is obtained for each sub-region, the process is called local thresholding. If the histogram of the image is assumed to have only one valley between two peaks, the technique is classified as bi-level thresholding, otherwise it is classified as multi-thresholding[94]. A very well known automatic threshold selection method for bi-level global thresholding is Otsu's method[123]. It provides satisfactory estimates when the number of pixels in the two classes is close to each other[94]. *Reddi et al.*[124] proposed a faster version of Otsu's method extended to multi-thresholding. A number of alternative automatic threshold selection algorithms have been developed over the years[94, 125].

An effective alternative to thresholding techniques is provided by mixture modeling and data clustering algorithms. An advantage of clustering approaches (albeit not so important for PET imaging) is that they can be applied not only to grayscale images, but also be used to segment multi-band images. Their main drawback is the necessity of prior selection of an appropriate number of clusters[94]. Clustering separates objects into groups whose members are similar in some way. This can be described as a process of finding structure in unlabeled data. This sets it apart from methods which use classifiers and require pre-segmented images called training data. Clustering algorithms are very often referred to as unsupervised methods as they must iterate between segmenting the image and characterizing properties of each cluster in order to compensate for the absence of training data and, in a sense, perform self training[94]. Examples of unsupervised segmentation methods are K-means clustering[126], C-means clustering[127], FCM (Fuzzy C-means) clustering[128] and the expectation-maximization (EM) algorithm[76]. These methods have shown very promising results as tumor delineation techniques in PET related studies; however more research is necessary for adequate algorithm initialization in order to guarantee reliable clustering results and a reduction in computational time[129]. FCM and C-means clustering are very popular in Nuclear Medicine[94, 129-131] and transmission image segmentation[132]. A major drawback in histogram based segmentation methods

such as thresholding or clustering is the lack of spatial information. In other words they do not take into account the intensity values of the surrounding pixels. Notable exceptions to this are methods based on 2D histograms which, to some degree, consider contextual information. Thresholding and clustering techniques work reasonably well on noise-free images, with slow spatial variation in intensity, and non-textured images[94]. Spatial characteristics of the image are an important piece of information for segmentation which is typically not taken into account in the above mentioned methods.

In contrast to histogram based segmentation methods, contextual information can be fully exploited in region growing approaches. For this type of problem, it is postulated that pixels which are close to each other have similar intensities. The procedure starts by selection of seed pixels with the goal of growing a uniform connected region from each seed, until stopping criteria are met[94]. The most important assumption in region growing approaches is that the regions have nearly constant gray levels. This makes them very sensitive to noise and texture, which renders them less reliable[94]. Region growing has been very successful in SPECT studies[94, 133, 134]. There are other techniques, such as Markov Random Field methods[135], which attempt to incorporate spatial correlations into the segmentation process[94]. Other examples of segmentation techniques include, but are not limited to methods using classifiers[94, 136-139], watershed methods[27, 84, 95, 140], edge detection (Sobel detector, Canny detector)[94], artificial neural networks (ANN, PCNN, MLP, RBFN)[94], deformable models (snake model)[141], atlas-guided approaches[94] and methods based on analysis of time activity curves (TACs)[32, 142, 143].

5.2.1 Image Segmentation in Dynamic PET Imaging

Several clustering algorithms have been proposed for tissue segmentation in dynamic PET related studies[32, 142-146]. In these approaches, dynamic PET images are first reconstructed from projection data using clinically accepted reconstruction algorithms and the time activity curves (TACs) of individual pixels are grouped into a predetermined number of clusters. *Kamasak et al.*[147] have used

an alternate approach whereby the clustering of the regions in dynamic PET images occurs directly with projection data, while at the same time estimating the TACs of each cluster. However, many institutions do not have direct access to raw sinograms, making implementation of this technique impractical for anyone other than PET manufacturers. The main drawback of clustering algorithms is sensitivity to noise and the need for determining the appropriate number of clusters[94]. Signals that arise in Nuclear Medicine (such as PET images) and their associated TACs are extremely noisy, and sensitivity to noise presents potential problems for accurate tumor volume delineation.

An alternative approach that allows use of the spatial information available in PET images is the combination of Region Growing (RG) and K-means cluster analysis. A hybrid technique using this approach has been proposed by *Kim et al.*[148]. In this approach, cluster analysis is used to extract temporal kinetic features and provide seeds for RG. The number of clusters is fixed to eight. This number presumably corresponds to the number of major structures (including the tumor) in the brain; however it renders application of the technique to other sites somewhat arbitrary. The performance of this technique directly depends on an Optimal Image Sampling Schedule (OISS)[149] along with a five-parameter FDG kinetic model (tracer dependent) used to reduce the dimensionality of the problem while increasing the SNR of individual time frames. This, in turn, dictates the development of clinical imaging protocols with OISS in mind at every institution interested in the algorithm, which may not be a feasible option. Furthermore, this algorithm was designed and validated around images reconstructed with Filtered Backprojection (FBP). Clinical use of FPB is diminishing (and is now practically extinct) due to the poor image quality associated with this reconstruction technique.

5.3 Advantages of Dynamic (4D) PET Imaging Based Tumor Volume Delineation

Conventional 3D PET images are constructed from a single time integration of activity concentration. Correct utilization of the functional information conveyed

by these images for radiotherapy purposes requires accurate geometric delineation of the boundaries between cancerous volumes and surrounding uninvolved tissues. The major difficulty encountered in attempting to properly segment conventional PET images is the great latitude in uptake values observed at any single point in time. Neighboring regions with different metabolic activity rates can, for example, exhibit large variances in perceived activity levels depending on the exact timing of image acquisition. This effect may act to either accentuate or obfuscate the boundary between cancerous and healthy tissue. Two such volumes, for example, might be imaged at a particular point in time when their activity levels are coincidently similar and may thus be falsely classified as the same tissue[32, 106, 150]. This makes geometric segmentation of conventional PET images problematic with regard to the accuracy required for radiotherapy treatment planning. 4D PET (dynamic PET) allows the addition of one more variable from the time domain, describing the temporal evolution of the form or function of tissue-specific biochemical properties over a specific time interval. Adjacent regions which might have similar activity levels in a static PET image can exhibit completely different temporal evolutionary courses[32, 106, 150]. These temporal variations in activity concentration are represented by measured time activity curves (TACs)[151]. Appropriate mathematical models may be used to describe the time activity curves (TACs) of a region or volume of interest, thus allowing the assessment of biological parameters. The shape of a TAC is indicative of tissue-specific biochemical properties, thus providing important insight into tracer kinetics and tissue uptake[32, 106, 152, 153]. For example, in studies with FDG, cancerous tissue exhibits steep upward TACs, while normal tissues exhibit flat TACs[32, 107, 154-157]. The use of TACs has the potential to significantly improve tumor volume delineation in PET imaging. In the following chapter we present a novel tumor volume delineation algorithm (SSRG/4D-PET) which exploits differences in the time activity curves (TAC) between tumor and surrounding background to more accurately delineate the geometric extent of PET defined tumor volumes for use in radiotherapy treatment planning.

6. Single Seed Region Growing Algorithm for Tumor Volume Delineation in Dynamic PET imaging (SSRG/4D-PET)[1]

The main goal of the research described in this chapter is the development of a partially-supervised tumor delineation algorithm incorporating both the temporal and spatial information available in dynamic PET studies for use in radiation therapy treatment planning. To this end, voxel TACs are analyzed and their differences exploited, without regard to preserving their "true" shape in the process. The performance of the proposed algorithm was evaluated first using both semi-empirical digital phantoms (homogeneous and heterogeneous) and a clinically-relevant phantom imaged on an Allegro-PET whole body PET scanner (Philips Medical Systems Inc.). Evaluation of the algorithm was also performed using clinical patient images acquired with a Philips Gemini GS PET/CT scanner (Philips Medical Systems Inc.).

6.1 Theory and Simulation Study

6.1.1 SSRG/4D-PET – The Algorithm

The SSRG/4D-PET algorithm was implemented in MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA). A partially-supervised (semi-automatic) approach was pursued in order to allow an expert reader to utilize the information available from other imaging modalities (CT or MR) routinely used in conjunction with PET (PET/CT or PET/MR). Two manual interventions are required: first, the definition of a mask around the structure of interest (e.g. the entire tumor volume, including reasonable margins) and second, the selection of a seed voxel within that structure. The subsequent delineation of the structure (tumor) within the mask is performed automatically by the algorithm, based on

[1] A version of this chapter has been published: A. R. Teymurazyan, R. S. Sloboda, T. R. Riauka, et al., "Single Seed Region Growing Algorithm in Dynamic PET Imaging (SSRG/4D-PET) for Tumor Volume Delineation in Radiotherapy Treatment Planning: Theory and Simulation," Nuclear Science, IEEE Transactions on 59 (5), 2020-2032 (2012).

comparison of the TAC of every voxel within this mask (candidate TAC) to the seed TAC.

PCA is employed in a pre-processing stage to produce a dynamic data set with a greatly reduced level of noise while enhancing contrast, thereby optimizing signal-to-noise ratio (SNR). This general approach has been successfully applied to PET data by other researchers[49, 82]. Denoising with PCA is completely independent of the kinetic model used and thus does not contain model-based restrictions. In dynamic PET data the maximum number of principal components corresponds to the number of sequence time frames. However, the majority of the total variance can be accounted for by the first two or three components. The remaining components are comprised mainly of noise and can be discarded without significant loss of useful information. The first principal component is formed as a linear combination of original variables containing the largest variance. Incidentally, the first principal component of a PET image volume will contain all high uptake regions (which often correspond to tumors) and can be expected to account for a significant amount of the total variance in the data. Since our intent is specifically the delineation of tumor volumes for the purpose of radiation treatment planning, only the first principal component is used in this study. It provides both the highest SNR and best discrimination between tumor and background tissue. In this work, dynamic PET data sets are processed by PCA slice-wise in the image domain. PCA of dynamic PET data requires the handling of large amounts of data. Slice-wise (as opposed to volumetric) application of PCA is less demanding with regards to computing power. The disadvantage of slice-wise PCA is possible inconsistencies within the resulting dynamic data set as parts of same structure in different slices are treated independently from each other. This does not pose a significant problem in the proposed schema as we are not attempting to extract kinetic parameters from TACs, but rather reduce the dimensionality of the problem while obtaining a data set in which the tumor and background are easily separated. This is achieved via fitting of the TAC of each voxel within a manually selected mask. The choice of the TAC model is dictated by the desire to fit the general shape of the TAC using

a very simple model independent of the specific tracer used (i.e. avoiding kinetic modeling and model-dependent restrictions). The denoising step prior to parameterization is necessary in order to eliminate local short term fluctuations in TACs which may significantly bias results. In the dynamic data set formed with first principal component images, all voxels (tumor and background) in the masked volume have the same kinetic characteristics. Only the amplitude parameter of the model function is therefore required for successful discrimination of tumor voxels. The parametric 3D image volume formed from the amplitude parameters is used in the region growing step of the algorithm. In the initial stage of the 3D seeded region growing process, a hard threshold separating regions of high uptake (large amplitude parameter) from regions of low uptake is automatically calculated. This calculated threshold value is used to compute the difference in amplitude between the threshold and the amplitude of the seed voxel selected in the high uptake region. This difference is used in the stopping condition governing voxel membership in the target structure. Any voxel within the mask satisfying the stopping condition is appended to the tumor volume if it is spatially connected to a voxel within the tumor region. The five main stages of the proposed tumor delineation algorithm for dynamic PET imaging are shown in **Figure 6.1**: 1) noise and artifact suppression; 2) selection of seed coordinates and a tumor encompassing mask of interest from the denoised 4D data; 3) fitting of candidate Time Activity Curves (TACs); 4) 3D-SSRG using a parametric image volume extracted from TACs; 5) tumor volume refinement.

6.1.1.1 Noise and Artifact Suppression

Prior to delineation, PCA is applied to the 4D Dynamic PET image set which is not corrected for radioactive decay. This is required by the TAC model function used and the algorithm will back-decay correct the input data if necessary. Each transverse slice comprising the 3D spatial component of the 4D image volume is processed separately by the PCA code. The temporal component of the 4D dynamic data set consists of n time frames, resulting in n pixels for every (*i, j*) coordinate point (pixel) in a given transverse slice. These pixels may be arranged

in the form of a column vector. For an M×N image (single transverse slice), there will be a total of $M \times N$ such n-dimensional vectors comprising all pixels in the n time frames[95]. The PCA code then operates on a two-dimensional feature matrix of size $n \times (M \times N)$, produced in unfolding the multidimensional image data matrix (M rows, N columns, and n variables). Slices reconstructed using only the first Principal Component (PC1) images are used to produce a noise and artifact suppressed dynamic image volume (4D; three spatial coordinates and one temporal coordinate)[49, 82, 95].

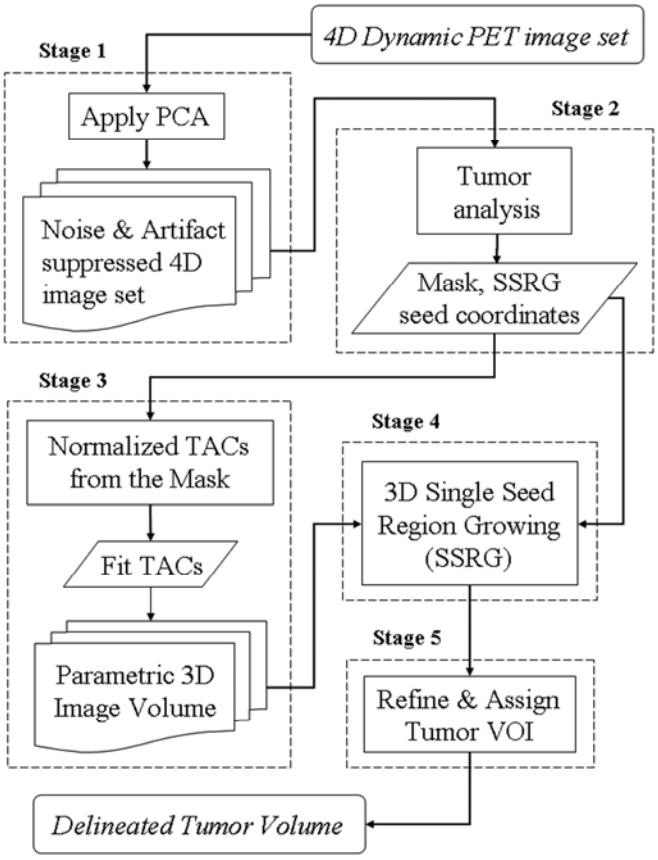

Figure 6.1: Flowchart of proposed semi-automatic algorithm.

6.1.1.2 Seed and Mask Selection

To minimize computational time, the user is able to manually select a 3D region around the target structure containing all voxels to be used for further analysis.

The noise and artifact suppressed dynamic data set produced in stage one is averaged in the temporal dimension, resulting in a 3D volume used in this step only. A clinical expert (physician) then defines a tumor encompassing mask (rectangular cuboid in the current implementation of our software) and also selects the seed site (single voxel), which is characterized as belonging to the target structure (tumor) with a high degree of confidence. The seed coordinates for 3D SSRG are thus extracted from the region of the highest activity level within the tumor encompassing volume.

6.1.1.3 3D Parametric Image Volume Definition via Fitting of Candidate TACs With an Appropriate Model Function

All voxel TACs within the mask volume are parameterized using a three parameter model given by equation **6.1**, simulating both nuclear (radioactive) decay and physiological tracer uptake and release:

$$T_{TAC}(k,t) = CN(k) \cdot e^{-\lambda_1(k)\cdot t} \cdot \left(1 - e^{-\lambda_2(k)\cdot t}\right) \qquad \textbf{6.1}$$

where $T_{TAC}(k,t)$ is the number of counts in the k^{th} voxel; $CN(k)$, $\lambda_1(k)$ and $\lambda_2(k)$ are corresponding model parameters (physical decay + uptake release and tracer uptake respectively).

A Levenberg-Marquardt method with line-search[158-160] is used to solve the nonlinear curve-fitting (data-fitting) problem in a least-squares sense. It was discovered empirically that solutions are achieved faster if noise and artifact suppressed 4D image data is normalized to its maximum value. This also simplifies the initialization of the data fitting algorithm.

The shape parameters λ_1 and λ_2 describe the dynamics of decay in the PC1 image set only and are not used in further processing. The amplitude (scaling) parameter $CN(k)$ corresponds to the k^{th} voxel and is used to produce a parametric image volume at the voxel level required in the next stage.

6.1.1.4 Stopping Parameter Selection and 3D-SSRG Based on Parametric Image Volume Extracted from TACs

In this stage, the candidate TACs are compared with the reference (seed) TAC by means of their amplitude parameter $CN(k)$ and the target volume is grown by way of a Single Seed Region Growing (3D-SSRG) algorithm. All voxels ($CN(k)$) of the parametric image volume produced in the previous stage are compared with $CN(seed)$. Voxels satisfying a "stopping condition" are classified as belonging to the target structure until the next stage, where connectivity to the seed region is tested.

The stopping condition for 3D-SSRG is defined as follows; the k^{th} voxel is appended to the tumor region if the Euclidian distance between amplitude parameters $CN(k)$ and $CN(seed)$ is less than or equal to the stopping parameter S.

$$\left|CN\left(k\right)-CN\left(seed\right)\right|\leq S \qquad \textbf{6.2}$$

The stopping parameter S is defined as the Euclidean distance in one dimension between the threshold value T and amplitude parameter $CN(seed)$, where the $CN(seed)$ is the amplitude parameter corresponding to the seed voxel.

$$\left|T-CN\left(seed\right)\right|=S \qquad \textbf{6.3}$$

The value of the threshold T is chosen based on a variation of the simple iterative approach first suggested by *Ridler et al.*[161]. The procedure is defined as follows: 1) an initial estimate of the threshold value T is chosen as the midpoint between minimum and maximum values of the amplitude parameter CN in the masked volume; 2) the parametric image volume is segmented using the initial estimate for the threshold, producing two groups of voxels (groups A with CN values larger or equal to T, and group B with CN values less than T); 3) a new threshold value (T) is defined as the average of CN_A (average of values in A) and CN_B (average of values in B); 4) the previous two steps are repeated in iterative fashion until the difference in threshold values between successive iterations is smaller than a predefined parameter P. The value of $P = 0.5$ used in this work was adopted from the literature[95] and allows one to obtain thresholds comparable to

those obtained using Otsu's method[123], without explicit binning of the data into histograms.

6.1.1.5 Tumor Volume Refinement & Assignment

In the refinement stage, all voxels in a tumor region are required to be connected in a specific way (e.g. neighborhoods for 3D tumor volume are 6, 18, or 26-connected). 26-connectivity was used throughout this work (a user modifiable feature in the algorithm). Connectivity within the tumor region is ensured through morphological reconstruction[162] of the image marker under the image mask with the specified connectivity, where the marker (seed volume; one at the seed coordinates and zero everywhere else) and the image mask (tumor volume; ones for the tumor region and zeros everywhere else) are binary 3D image volumes of the same size. Lastly, the refined tumor volume is remapped into the original image volume (**Figure 6.2**).

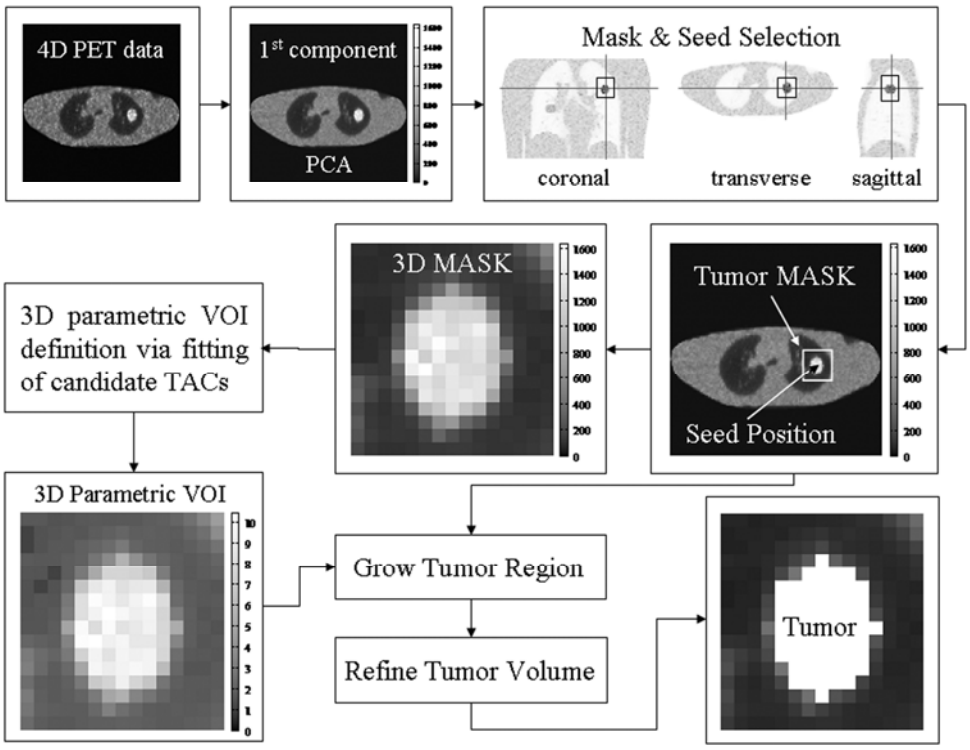

Figure 6.2: Workflow of proposed algorithm (SSRG/4D-PET)

The workflow (**Figure 6.2**) of SSRG/4D-PET can be summarized as follows: PCA is employed in a pre-processing stage to provide a noise and artifact suppressed 4D image set; seed coordinates are selected and a tumor-encompassing mask is defined; the data within this mask is then used to produce a parametric image volume by fitting each voxel's TAC (the candidate TACs) to a model function; the amplitude parameters of each candidate voxel are compared with the amplitude parameter of the seed voxel which serves as a reference value; thus the 3D tumor region is grown from a single seed by means of a seeded region growing approach. Finally, the refined tumor volume is remapped into the original image volume.

6.1.2 Algorithm Performance Evaluation Methods

To analyze algorithm performance, target delineation results obtained using our algorithm were compared to the known (true) phantom geometries. Two different performance (similarity) metrics were used: misclassification rate (R_m) and Dice coefficient (Dice). For each delineation, the number of voxels wrongly classified as belonging to a given structure (false positive (*FP*)), wrongly classified as not belonging to a structure (false negative (*FN*)), correctly classified as belonging to a structure (true positive (*TP*)), and correctly classified as not belonging to a structure (true negative (*TN*)) were tallied and used to compute these similarity metrics.

The misclassification rate R_m, is defined as:

$$R_m = \frac{FN + FP}{TP + FN} \times 100\% \qquad \textbf{6.4}$$

and is the total percentage of voxels incorrectly assigned by the delineation algorithm in question (*FN*+*FP*), compared to the total number of voxels in the "true" target volume (*TP*+*FN*). The misclassification rate is an overlap error measure (a value of 0 indicates perfect agreement). In contrast, the Dice coefficient (also known as Sorensen-Dice coefficient) is an overlap agreement

measure. For this metric, a value of zero indicates data sets with no common members and value of one indicates perfect agreement.

The Dice coefficient is defined as follows,

$$Dice = \frac{2 \times TP}{(FP + TP) + (TP + FN)} \qquad \textbf{6.5}$$

6.1.3 Simulation Study: Homogeneous Phantom

To test the validity of the SSRG/4D-PET delineation algorithm, a semi-empirical digital phantom was created to simulate fluorine-18-fluorodeoxyglucouse ([^{18}F]-FDG or FDG) metabolism in 7 clinically relevant tissue types. These tissues include three normal structures: skeletal muscle (SM), lung (L), and myocardium (M); two malignant lesion types: Stage III non-small-cell lung cancer (M1), and lung carcinoma (M2); and two benign lesion types: lung aspergillosis (B1), and lung coccidomycosis (B2). Geometric definition of the body contour and the three normal internal structures was achieved by segmentation of a representative anatomically correct 3D CT data set. Lesions were added within the lung as clinically relevant sized[163] spheroid and ellipsoid volumes (see **Table 6.1**).

Table 6.1: FDG two-tissue compartmental model parameters

Compartment (or Pathology)	K_1 (ml/min/g)	k_2 (min^{-1})	k_3 (min^{-1})	k_4 (min^{-1})	Size (cm)	Size (pixels)
Myocardium (M)	0.196	1.022	0.149	0.010	-	-
Lung (L)	0.014	0.291	0.006	0.000	-	-
Skeletal Muscle (SM)	0.045	0.163	0.077	0.000	-	-
Stage III non-small-cell lung cancer (M1)	0.084	0.021	0.072	0.000	3.2×3.2	8×8
Lung Carcinoma (M2)	0.139	0.296	0.164	0.000	3.2×4.0	8×10
Lung Aspergillosis (B1)	0.1993	0.9778	0.240	0.010	4.0×4.0	10×10
Lung Coccidomycosis (B2)	0.181	0.8692	0.039	0.001	2.4×4.0	6×10

Post delineation, the original CT data set was downsampled to yield a 3D image consisting of 57 slices each with a matrix size of 128×128 and cubic voxels with 4 mm sides. Generation and visualization of this digital phantom was performed with MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA). FDG was chosen as the radiopharmaceutical tracer to be modeled as it is used in over 90% of all PET procedures in current clinical practice.

Table 6.2: Time sampling schemes used

Index (TSS_i)	Sampling Schedule	Number of frames × frame duration (in seconds)
1	157 min, 31 frames	31×300
2	25 min, 19 frames	10×12, 2×30, 2×60, 1×90, 1×210, 2×300, 1×600
3	55 min, 20 frames	10×12, 2×30, 2×60, 1×90, 1×210, 2×300, 1×600, 1×1800
4	115 min, 22 frames	10×12, 2×30, 2×60, 1×90, 1×210, 2×300, 1×600, 3×1800
5	33 min, 18 frames	12×10, 2×40, 2×300, 2×600
6	73 min, 22 frames	12×10, 2×40, 2×300, 6×600
7	31 min, 5 frames	2×41, 1×136, 1×567, 1×1104
8	60 min, 5 frames	1×41, 1×136, 1×567, 1×1145, 1×1711
9	30 min, 18 frames	8×15, 6×30, 3×300, 1×600
10	30 min, 19 frames	8×15, 6×30, 5×300
11	60 min, 21 frames	8×15, 6×30, 3×300, 4×600
12	60 min, 23 frames	8×15, 6×30, 7×300, 2×600

6.1.3.1 Homogeneous Phantom Generation

The kinetics of FDG uptake in each tissue type was simulated using a two-tissue compartmental model[164]. This model consists of a compartment representing arterial plasma and two tissue compartments: free (unmetabolized) FDG and its intra-cellularly trapped metabolite FDG-6-phosphate (FDG-6-PO_4). If the plasma time activity curve ($P_{TAC}(t)$, also referred to as the input function) and the relevant model parameters are known, then for the case of homogeneous tissue uptake, the PET activity concentration ($C_{TAC}(t)$) in tissue can be calculated as follows,

$$C_{TAC}{}^{[m]}(t) = \frac{K_1^{[\mathrm{m}]}}{\alpha_2^{[\mathrm{m}]} - \alpha_1^{[\mathrm{m}]}} \times \left[\begin{array}{l} \left(k_3^{[\mathrm{m}]} + k_4^{[\mathrm{m}]} - \alpha_1^{[\mathrm{m}]}\right) \cdot e^{-\alpha_1^{[\mathrm{m}]} \cdot t} + \\ \left(\alpha_2^{[\mathrm{m}]} - k_3^{[\mathrm{m}]} - k_4^{[\mathrm{m}]}\right) \cdot e^{-\alpha_2^{[\mathrm{m}]} \cdot t} \end{array} \right] \otimes P_{TAC}(t) \qquad \textbf{6.6}$$

where the symbol ⊗ represents convolution; K_1 is the perfusion (ml/min/g) and the other rate constants (k_2, k_3 and k_4) are in units of min^{-1}. Constants α_1 and α_2 are functions of the rate constants and index m indexes the tissue type simulated.

The kinetic parameters associated with each tissue type are derived from the literature[165-167] and summarized in **Table 6.1**. The plasma time activity curve, $P_{TAC}(t)$, which serves as the input function to this two-tissue compartmental model, may be obtained empirically by invasive sampling of arterial blood [168]. Alternatively, a semi-empirical approach may be taken whereby $P_{TAC}(t)$ is determined analytically according to published clinical data [168, 169]. The latter approach was adopted for this work according to the methodology proposed by *Feng et al.*[170]. The plasma time activity curve was calculated using the expression,

$$P_{TAC}(t) = \begin{cases} \left[A_1(t-\tau) - A_2 - A_3\right] e^{\lambda_1 (t-\tau)} \\ + A_2 e^{\lambda_2 (t-\tau)} + A_3 e^{\lambda_3 (t-\tau)} t & t > \tau \\ 0 & t \le \tau \end{cases} \qquad \textbf{6.7}$$

where the delay factor $\tau = 0.537$, and $A_1 = 5.987$, $A_2 = 0.0453$, $A_3 = 0.0881$, $\lambda_1 = -6.6887$, $\lambda_2 = -0.2458$, and $\lambda_3 = -0.0181$ (*Feng et al.* [169]).

PET measurements of tracer concentration ($\tilde{C}_{TAC}(t_n)$) are acquired over finite time intervals and are weighted time averages over the collection interval (Δt_n) of a scan time frame,

$$\tilde{C}_{TAC}^{[m]}(t_n) = \frac{1}{\Delta t_n} \int_{t_{n-1}}^{t_{n-1}+\Delta t} C_{TAC}^{[m]}(t) dt \qquad \textbf{6.8}$$

where $t_n = t_{n-1} + \Delta t_n$ and t_{n-1} is the start time of n^{th} frame. The scanning interval Δt_n of the n^{th} time frame is defined by the time sampling schedule. The nominal imaging interval for FDG-PET is 60 min. It has been shown, however, that for

lung cancer, FDG kinetic parameters derived from time intervals as short as 30 minutes can provide comparable results[163]. Twelve time sampling schedules (see **Table 6.2**) adopted from the literature[149, 163, 168, 169] were used to generate simulated dynamic PET data sets and aid in evaluating algorithm performance.

Tracer concentrations were assigned to each voxel in every time frame using the kinetic parameters of each tissue type as per **Table 6.1**. Finally, noise was added to these data sets. Gaussian noise was created using a random number generator and was added to the noiseless tracer concentrations $\tilde{C}_{TAC}^{[m]}(t_n)$ accumulated in each voxel representing the m^{th} tissue type of the n^{th} time frame of a dynamic PET scan as follows,

$$\tilde{C}_{TAC}^{[m]}(t_n)^N = \tilde{C}_{TAC}^{[m]}(t_n) + \sigma^{[m]}(t_n) \times G_N(0,1) \qquad \textbf{6.9}$$

where $\tilde{C}_{TAC}^{[m]}(t_n)^N$ is the noisy activity concentration of the n^{th} time frame, $\sigma^{[m]}(t_n)$ is the standard deviation of noise, and $G_N(0,1)$ is a random number drawn from a Normal distribution with a mean of zero and a standard deviation of one. The standard deviation $\sigma^{[m]}(t_n)$ is given by (**6.10)** as follows[165],

$$\sigma^{[m]}(t_n) = \left(\frac{\alpha \times \tilde{C}_{TAC}^{[m]}(t_n)}{\Delta t} \right)^{1/2} \qquad \textbf{6.10}$$

where the multiplier α = {0.1, 0.5, 1.0, 2.0, 4.0} serves to define five different noise levels adopted from the literature[168].

The final step in simulating realistic dynamic PET measurements was the forward projection (Radon transform) of these activity concentration distributions to produce temporally sequenced sinograms. This projection data was then reconstructed using Maximum Likelihood Expectation Maximization (ML-EM) [59]. This was performed using the *Black Box Toolbox*[171] implementation of this algorithm. Throughout, unless otherwise stated, simulations were performed with a noise level of $\alpha = 0.1$ and a time delay of 2 minutes between tracer injection and simulated data collection.

6.1.3.2 Homogeneous Phantom: Results and Discussion

Dice coefficients for the four lesion types are plotted as a function of TSS index (see first column in **Table 6.2**) in **Figure 6.3**. One may readily observe that the Dice coefficients for the malignant lesions (M1 and M2) are very close to unity with negligible variation for TSS_2 through TSS_{11} (average of 0.990 and standard deviation of 0.003), indicating very good segmentation performance of our algorithm for this lesion. Diminished algorithm performance is realized for TSS_1 (0.937 for M1 and 0.877 for M2) and TSS_{12} (0.791 for M1 and 0.830 for M2), giving evidence to the fact that not all TSS schemes produce similar results and no technique is universal in scope and capability.

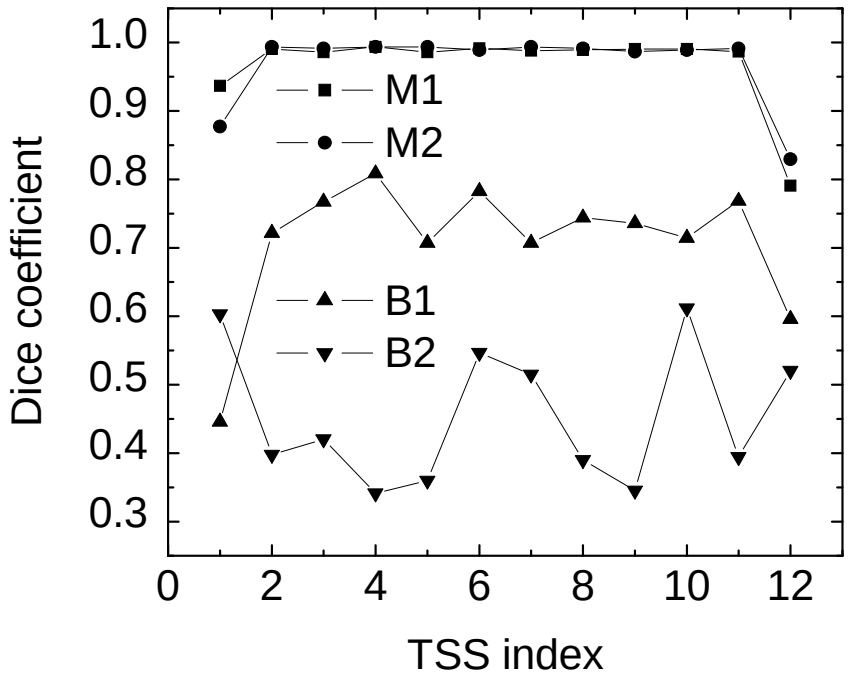

Figure 6.3: Dice coefficient as function of time sampling schedule (TSS) for the four different types of lesions ($\alpha = 0.1$, delay = 2 min). Delineation algorithm used: SSRG/4D-PET.

A similar trend, albeit with considerably smaller Dice coefficients and much greater variability, is obtained for benign lesion B1. Here, the best performance is realized for TSS_4 with a Dice coefficient of 0.809. Least effective results are obtained for benign lesion B2 with Dice coefficients as low as 0.342 and 0.345 realized for TSS_4 and TSS_9 respectively.

Mean values of the Dice coefficient and misclassification rate (R_m), averaged over the twelve TSS are tabulated in **Table 6.3**, along with their corresponding standard deviations (STD). All simulations were performed with a noise level of $\alpha = 0.1$ and a time delay of 2 minutes.

Table 6.3: Mean values of *Dice* and R_m measures for the SSRG/4D-PET target delineation algorithm averaged across the twelve time sampling schedules (**Table 6.2**). Phantoms were generated with $\alpha = 0.1$ and delay = 2 min

Pathology	Dice coefficient		R_m (%)	
	Average	STD	Average	STD
Stage III non-small-cell lung cancer (M1)	0.97	0.06	7.02	14.06
Lung Carcinoma (M2)	0.97	0.05	6.70	12.27
Lung Aspergillosis (B1)	0.71	0.10	83.55	41.34
Lung Coccidomycosis (B2)	0.45	0.10	129.54	67.17

Values of Dice coefficient very close to unity for the malignant lesions (M1 and M2) indicate very few false positives and false negatives, corresponding to higher sensitivity (capability of the algorithm to recognize a lesion) and specificity (capability of the algorithm to distinguish a lesion from normal tissue)

when compared to the results obtained for benign (B1 and B2) lesions. These results can be attributed to the shape of the TACs involved, which are indicative of tissue-specific biochemical properties[32]. For example, in studies with FDG, cancerous tissue TACs continually rise for an extended period of time post injection (up to 60 min and more in some studies[165]), while TACs of normal tissues quickly reach a shallow maximum followed by a gradual wash-out[32, 156, 165]. The algorithm employed here relies on the presence of a sufficient difference between the target and surrounding background TACs.

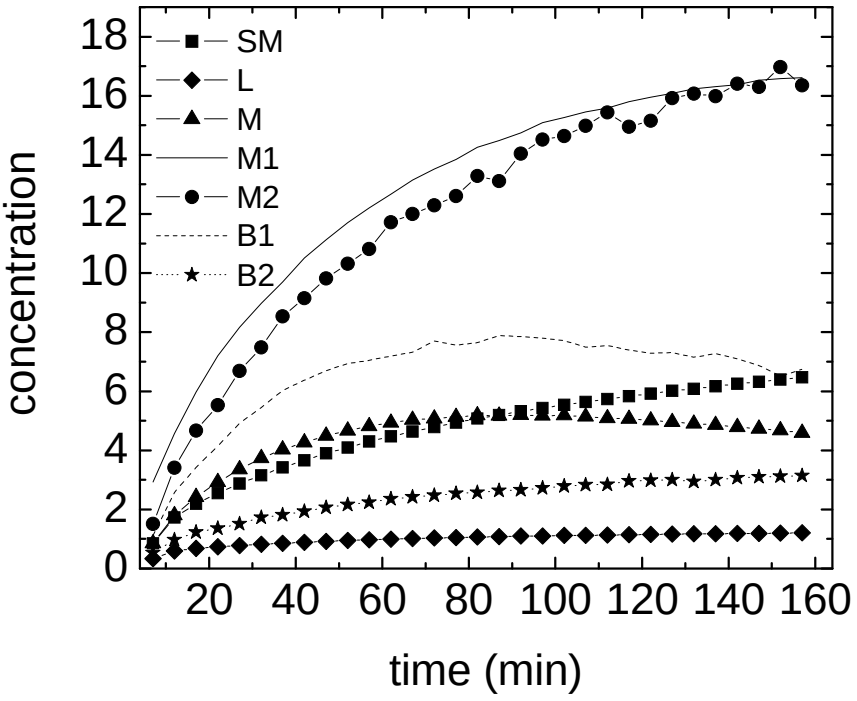

Figure 6.4: TACs (average values for each compartment) assigned to the structures in the mathematical simulation according to TSS_1 (31 frames × 5 min, $\alpha = 0.1$, delay = 2 min)

Figure 6.4 shows that the TACs and the concentration levels for B1 and B2 are similar to that of normal tissues (Myocardium (M), Lung (L), and Skeletal

Muscle (SM)), while M1 and M2 display steep upward TACs characteristic of tumors. B1 exhibits steeper uptake and larger concentrations relative to B2 and to surrounding normal tissue, leading to the better agreement between delineated and true volume for B1 as compared to B2. These results suggest caution when applying the algorithm to lesions with TACs insufficiently distinct from normal, healthy tissue.

To investigate the effect of noise on the performance of SSRG/4D-PET at different noise levels, data sets with five noise levels (α = {0.1, 0.5, 1.0, 2.0, 4.0} [168]) were generated for TSS_4. Comparison of the delineation of M1, M2, B1 and B2 at these noise levels is shown in **Figure 6.5**.

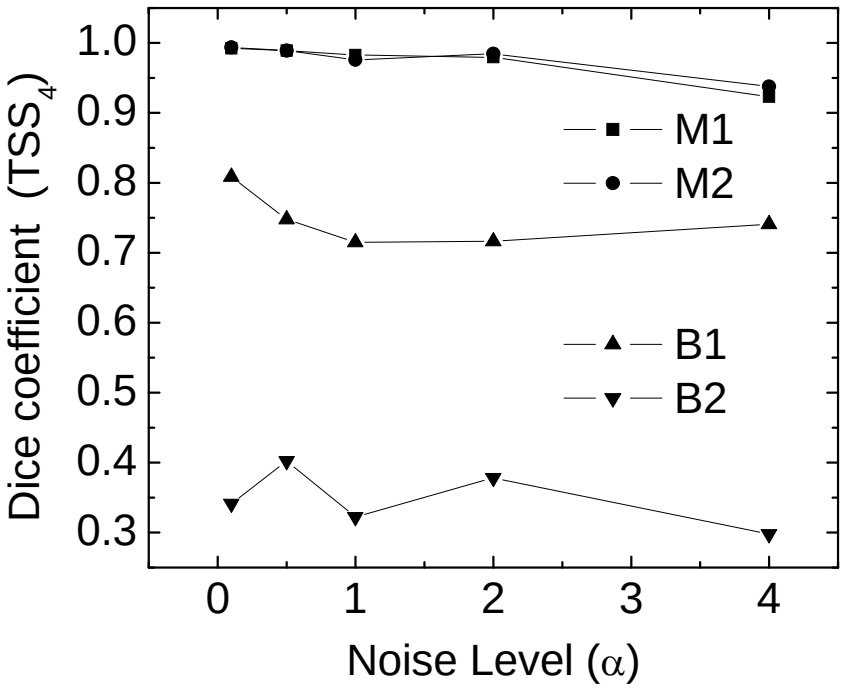

Figure 6.5: Dice coefficient comparing ground truth to simulated tumors for the dynamic phantom generated with TSS_4 at five different noise levels (α = {0.1, 0.5, 1.0, 2.0, 4.0})

While an almost linear decrease in algorithm (SSRG/4D-PET) performance is observed with increasing noise levels for malignant lesions, no clear trend can be discerned for benign lesions due to larger variations in Dice coefficient values.

To investigate the effect of a delay between the time of injection and start of the PET scan on algorithm performance, five (*delay* = {2, 5, 15, 30, 60} min) different time delay intervals were used to generate simulated dynamic PET data sets according to TSS_1.

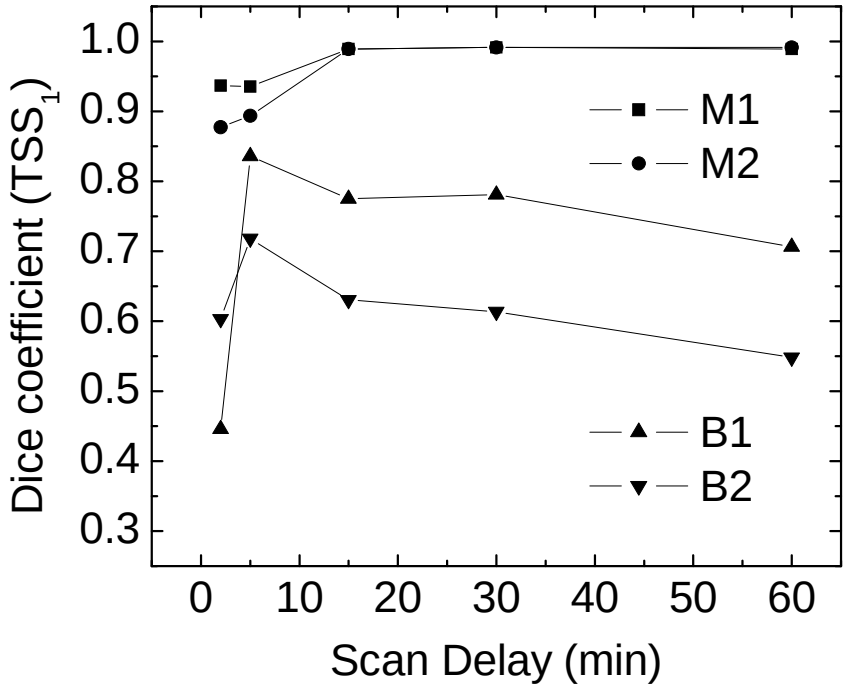

Figure 6.6: Dice coefficient comparing ground truth to simulated tumors for the dynamic phantom generated with TSS_1 for five different time intervals between the time of injection and start of PET scan (delay = {2, 5, 15, 30, 60} min)

The analysis of the effect of a delay between the time of injection and start of a PET scan on algorithm performance (**Figure 6.6**) reveals a significant

improvement in delineation for both malignant lesions (M1 and M2), when increasing the time delay from 2 to 15 minutes as Dice values increase from 0.937 to 0.989 for M1 and from 0.877 to 0.989 for M2. This steep increase observed for intervals up to 15 min is followed by a plateau region (Dice is equal to 0.99 for both lesions). For the benign lesions (B1 and B2), a dramatic increase in performance is observed in going from a 2 to a 5 min interval (0.45 to 0.84 for B1 and 0.60 to 0.72 for M2) followed by a systematic decrease in Dice values for time delays of greater than 5 minutes. Clearly, of those examined, a 5 minute time delay yields optimum results for both B1 and B2.

To test the robustness of the proposed delineation technique based on single seed region growing with respect to seed location selection, we randomly selected 30 seed locations within lesion M1. For each randomly selected seed location as shown in **Figure 6.7a**, delineation of the M1 lesion was performed for four TSS (TSS1, TSS4, TSS5, and TSS6) at $\alpha = 0.1$ and a 2 min time delay. Using Dice coefficients the resulting M1-delineations were compared with the delineations based on the original manual seed selection (**Figure 6.3**).

The Dice coefficient as a function of stopping parameter S for the resulting delineations is shown in **Figure 6.7b**. For 22 out of 30 seed locations with stopping parameter above 0.25 for all four TSS, the Dice coefficient is nearly constant with the average value of 0.97 and standard deviation of 0.03, consistent with the average Dice coefficient for delineation based on a manually selected seed (0.97 with standard deviation 0.06, **Table 6.3**). For the remaining eight seed locations, the Dice coefficient is decreasing with decreasing stopping parameter. Examination of the spatial distribution of these seed locations reveals that seed locations with $S < 0.25$ are located on the periphery of the tumor (open symbols in **Figure 6.7a**). It should be noted, however, that seeds in the border regions are not likely to be selected by a user in a clinical setting due to partial volume effects.

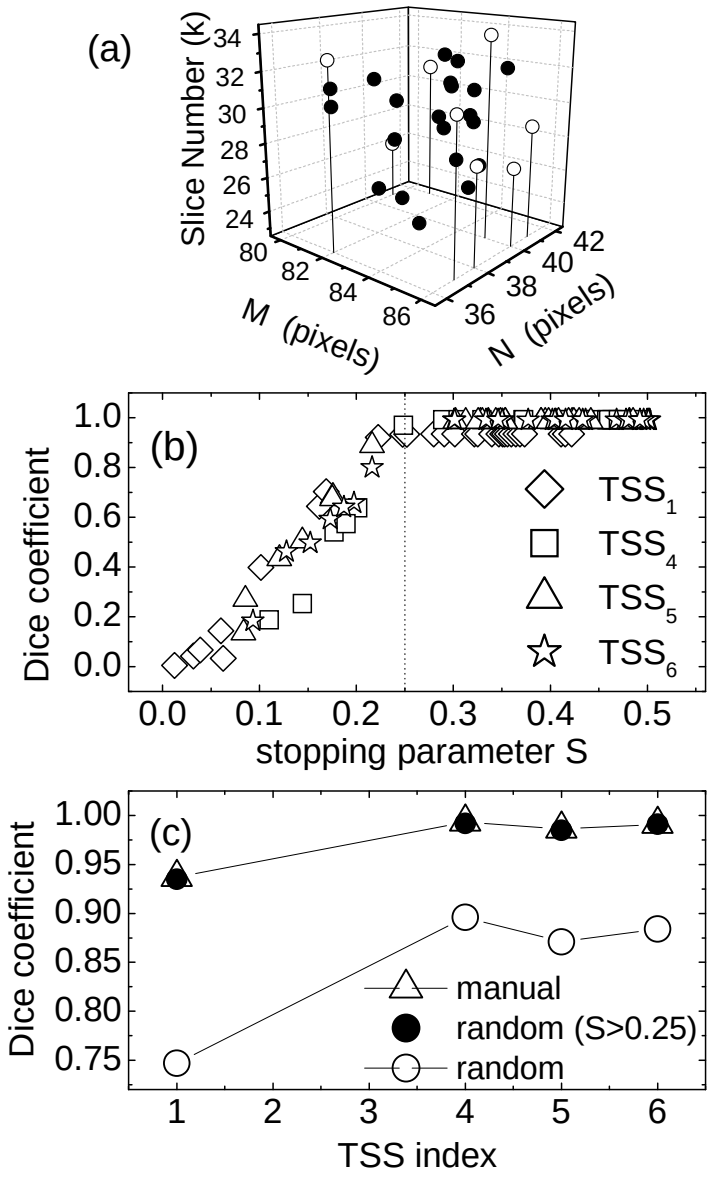

Figure 6.7: Impact of random seed selection on SSRG/4D-PET algorithm performance: (a) the 3D spatial distribution of randomly selected seeds within tumor volume; Dice coefficient as a function of stopping parameter S (b) and TSS as indexed in **Table 6.2** (c)

Figure 6.7c depicts the comparison of average Dice coefficient for each of the four TSS for delineation based on a seed manually selected by the user to delineation based on seeds chosen randomly. The values of Dice coefficients averaged for all 30 seeds (line + open circle) are nearly 10 % lower compared to values obtained with manual seed selection (line + triangle). However, when the eight points located on the periphery of the tumor are eliminated, the agreement between the randomly selected and user-selected seed Dice coefficients is excellent. These results suggest that the single seed selection method may be used successfully in the proposed algorithm (SSRG/4D-PET) without inducing significant bias, provided that seed locations in peripheral regions are avoided.

Finally, four alternate approaches for target (M2 and B2) delineation are compared to the algorithm proposed here: an algorithm based on K-means clustering with the number of clusters fixed to two, a fixed (40 %) threshold of maximum signal intensity, global automatic image thresholding[95] using Otsu's method[123] and adaptive (local) automatic thresholding[95, 123]. K-means clustering was performed using MATLAB function "k-means" (Statistics Toolbox; Version 5.0.1) and the L2 norm as an index of similarity. Adaptive thresholding was implemented by taking the morphological opening of an image, then applying a global image threshold (obtained using Otsu's[123] method) to the result. A square 8×8 pixels structuring element was used in morphological operations in adaptive thresholding. Tumor volumes were delineated on the same masked volume used in the SSRG/4D-PET algorithm. Prior to delineation, the dynamic data set was averaged in the temporal direction resulting in a 3D summation image volume over all time frames. Delineation results were evaluated by Dice coefficient and misclassification rate (R_m) for two different lesion types (M2 and B2), applied directly to the 3D delineation results. Data based on the performance of these five different algorithms is presented in **Table 6.4** and

Table 6.5. In comparison to other algorithms, SSRG/4D-PET yields the highest Dice value and the lowest misclassification rate for both lesion types.

Table 6.4: Comparison of delineation of the M2 lesion using five different algorithms. Values of Dice and R_m measures were averaged across twelve time sampling schedules (**Table 6.2**). Phantoms were generated with $\alpha = 0.1$ and delay = 2 min

Method	Dice coefficient		R_m (%)	
	Average	**STD**	**Average**	**STD**
SSRG/4D-PET	0.97	0.05	6.70	12.27
K-means clustering with 2 clusters	0.91	0.14	27.84	58.81
Fixed threshold of maximum signal intensity (40%)	0.90	0.19	44.87	124.93
global automatic thresholding (Otsu[123])	0.78	0.28	106.55	162.44
adaptive (local) automatic thresholding (Otsu[123])	0.90	0.16	33.01	74.29

Table 6.5: Comparison of delineation of the B2 lesion using five different algorithms. Values of Dice and R_m measures were averaged across twelve time sampling schedules (**Table 6.2**). Phantoms were generated with $\alpha = 0.1$ and delay = 2 min

Method	Dice coefficient		R_m (%)	
	Average	**STD**	**Average**	**STD**
SSRG/4D-PET	0.45	0.10	129.73	67.17
K-means clustering with 2 clusters	0.22	0.06	256.94	39.86
Fixed threshold of maximum signal intensity (40%)	0.10	0.04	227.39	28.75
global automatic thresholding (Otsu[123])	0.19	0.07	257.89	84.31
adaptive (local) automatic thresholding (Otsu[123])	0.15	0.13	212.10	28.04

6.1.4 Simulation Study: Heterogeneous Phantom

To evaluate the performance of the SSRG/4D-PET delineation algorithm in the presence of heterogeneous uptake, a semi-empirical digital phantom simulating FDG metabolism in three clinically relevant tissue types was generated. These tissues include two normal structures: Skeletal Muscle (SM) and Liver (L); and one malignant lesion: Liver Tumor (LT). Geometric definition of the body contour, the two normal internal structures and the tumor was achieved by manual segmentation of a representative fused 3D PET/CT data set. Segmentation was performed using the Region Of interest Visualization, Evaluation, and image

Registration (ROVER) software package (ABX GmbH, Germany). The resulting segmented 3D image volume consisted of 45 slices, each a matrix of 144×144 cubic voxels with 4 mm sides. Generation and visualization of this digital phantom was performed with MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA).

6.1.4.1 Heterogeneous Phantom Generation

Heterogeneous tissue simulations were performed using an FDG kinetic model developed by *Wu et al.*[172]. This is an extension of the "homogeneous" compartmental model given by (**6.6**), modified to account for tissue heterogeneity effects and refined for FDG-PET kinetic studies in liver metastases. According to this model, PET activity concentration ($C_{TAC}^{normal}(t)$) in normal tissues (Skeletal Muscle or Liver) can be calculated as a sum of activity concentration in normal cells (SM or L) and in a vascular compartment,

$$C_{TAC}^{normal}(t) = C_{TAC}^{normal\,cell}(t) + v_2 P_{TAC}(t) \qquad \textbf{6.11}$$

where v_2 is the vascular compartment volume (fractional volume) accounting for the vascular space in normal tissues (SM or L), $C_{TAC}^{normal\,cell}(t)$ is the homogeneous uptake in the normal tissue ($C_{TAC}^{L\,cell}(t)$ or $C_{TAC}^{SM\,cell}(t)$) given by (**6.6**), and $P_{TAC}(t)$ is the plasma time activity curve given by (**6.7**). Substitution of tissue blood radioactivity with the plasma radioactivity ($P_{TAC}(t)$) has been validated by *Phelps et al.*[173]. Each voxel within the tumor volume is assumed to be comprised of both a normal (L) and tumor (LT) cell population. Each population is assumed to have a homogenous compartment behavior given by (**6.6**) and its own characteristic rate constants (K_1, k_2, k_3 and k_4). Thus,

$$C_{TAC}^{LT}(t) = w_1 C_{TAC}^{LT\,cell}(t) + w_2 C_{TAC}^{L\,cell}(t) + v_1 P_{TAC}(t) \qquad \textbf{6.12}$$

where w_1, and w_2 are the percent mass weights of the liver tumor (LT) and liver (L) normal cell population respectively ($w_1 + w_2 = 1$), and v_1 is the fractional vascular compartment volume within the tumor. The liver cell population ($C_{TAC}^{L\,cell}(t)$) in the tumor is assumed to have the same characteristic rate constants

as a normal liver cell population. Rate constants for liver tissue (K_1 = 0.864 ml/min/g, k_2 = 0.981 min^{-1}, k_3 = 0.005 min^{-1}, k_4 = 0.160 min^{-1}; mean values from 10 volunteers) and tumor (K_1 = 0.243 ml/min/g, k_2 = 0.780 min^{-1}, k_3 = 0.101 min^{-1}, k_4 = 0.000 min^{-1}; tumor source: liver metastases, mean values from 25 melanoma studies) were adopted from a study by *Wu et al.*[172]. For skeletal muscle tissue, values presented in **Table 6.1** were used. In this simulation, the tumor (LT) cell population fraction w_1 in each voxel of the tumor volume was derived from the same representative 3D PET/CT data set which was used for geometric definition of the body contour, the two normal internal structures and the liver tumor contour. The intensity distribution within the liver tumor contour outlined on this clinical PET data was normalized to the maximum intensity value in the tumor. The weight distribution obtained was used in simulating heterogeneous tumor tissue. The simulated tumor was irregularly shaped with total volume of 8.1 cc. A ten percent contribution (v_1 = 0.1) of plasma activity was added to tumor concentrations, while twenty five percent (v_2 = 0.25) was added to the normal tissue TACs. For tumor and normal tissues, noiseless PET measurements of tracer concentration over finite time intervals are calculated as weighted time averages over the collection interval of a scan time frame as given by (**6.8**). A dynamic sequence of 28 frames (10×30 s, 5×60 s, 5×120 s, and 8×300 s) was simulated according to a time sampling schedule used in a study by *Janssen et al*.[32]. Finally, Gaussian noise comparable to the level in clinical PET studies was added to these data sets following (**6.9**), where the standard deviation is calculated from the noiseless tracer concentrations using (**6.10**). Noise levels in the last 40 min of tumor tissue, liver tissue and skeletal muscle tissue TACs were 4.9%, 3.4% and 3.3%, respectively[172]. To simulate the resolution blur characteristic of PET data, images were convolved with a PSF (Point Spread Function) representative of a modern PET system. The PSF was approximated by a Gaussian kernel, with an isotropic FWHM of 6 mm. The value of the FWHM is larger than the width of the nominal PSF (4.31 mm for the Philips Allegro/Gemini cameras in case of ^{18}F) attributable to the PET system and reconstruction method, in order to account for the additional blur caused by scatter[174].

6.1.4.2 Heterogeneous Phantom: Results and Discussion

In this evaluation, in addition to the three thresholding techniques and static K-means clustering method discussed earlier, the performance of the SSRG/4D-PET algorithm is compared to a dynamic tumor volume delineation technique proposed by *Janssen et al*.[32] (henceforth Dynamic K-means clustering algorithm). In the study by *Janssen et al*.[32] a dynamic sequence of 28 frames acquired according to a schedule of 10×30 s, 5×60 s, 5×120 s, and 8×300 s (number of frames × frame duration (in seconds)) is assumed. This time sampling schedule was also used in the heterogeneous phantom simulation in order to facilitate a comparison with the dynamic K-means clustering algorithm. The major constraints in the algorithm are: a) the slope of the TAC (reflecting the amount and rate of FDG uptake) is higher within the tumor tissue than within surrounding normal tissue; b) the number of clusters is set to 2 a priori. Prior to tumor delineation, several pre-processing steps are applied: 1) the volume of interest (VOI) is manually selected around the tumor to minimize the computational time and restrict the number of normal structures in close proximity to the tumor; 2) an edge-preserving bilateral filter[174] is applied to the data to attenuate the statistical noise and prepare the image for the de-blurring step; 3) to correct for blurring due to the point-spread function of the PET system, the data is de-blurred by means of Landweber's iterative algorithm[174]; 4) filtering with a moving average filter in the time domain is used to smooth short-term fluctuations and highlight longer-term trends within the TACs. In the tumor delineation stage, the TAC slope value is calculated for each voxel within the selected VOI. Slopes are calculated over the last eight time frames of the dynamic PET scan as the average of the sum of the differences in FDG activity between two (time) adjacent measurements. Finally, slope values of the TACs are clustered using the k-means clustering algorithm, with the number set to 2. The cluster containing voxels with the highest mean TAC slope is classified as tumor. We would like to stress for the reader that in this comparison we have used our implementation of the method proposed by *Janssen et al*.[32]. The algorithm was implemented with MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA). Because of the absence of

parameter values for the various filtering methods used in the publication by *Janssen et al*.[32], parameters from the referenced original work[174] were used for bilateral filtering and iterative de-blurring. Filtering in the time domain was performed using a moving average filter with window size spanning 3 samples. K-means clustering was performed using the MATLAB function "k-means" and an L2 norm as the index of similarity.

Since this dynamic K-means clustering algorithm utilizes only the last eight time frames of the 4D tumor volume, all other methods (including SSRG/4D-PET) evaluated here were also applied only to the last eight time frames. Prior to delineation, in the case of static methods (thresholding and K-means clustering), the dynamic data set was averaged in the temporal direction over the last eight time frames. Tumor volumes were delineated on the same masked volume used in the proposed algorithm (SSRG/4D-PET).

Table 6.6: Heterogeneous phantom study – comparison of delineation results for six algorithms

Method		**Dice coefficient**	**R_m (%)**
SSRG/4D-PET		0.84	27.56
Dynamic K-means clustering (*Janssen et al.*[32])	2 clusters	0.74	70.10
	3 clusters	0.70	68.50
	4 clusters	0.56	65.40
K-means clustering	2 clusters	0.55	165.40
	3 clusters	0.65	52.00
	4 clusters	0.64	52.80
Fixed threshold of maximum signal intensity (40%)		0.04	5468.50
global automatic thresholding (Otsu[123])		0.55	165.35
adaptive (local) automatic thresholding (Otsu[123])		0.57	148.82

Table 6.6 presents the segmentation results of these six different approaches evaluated according to Dice coefficient and misclassification rate (R_m). For completeness, we have also included the delineation results for both clustering methods (dynamic and static) with the number of clusters set to 3 and 4.

In clinical situations, depending on tumor size and location, it may prove difficult to select a VOI (Volume of Interest) around the tumor containing a single type of normal tissue. For this reason, since the simulated tumor was located in the left section of the superior sub-segment of the anterior portion of the liver, the manually selected VOI around the tumor contained both liver tissue and a small amount of simulated skeletal muscle tissue. This fact contributed to the high misclassification rates (high number of false positives) presented in **Table 6.6**. However it is only partially responsible for the complete failure of the fixed (40% of max) thresholding technique. The size and shape of the simulated tumor (and associated partial volume effects) are to blame. For small tumors a high threshold value (>50%) is more appropriate[33]. However, the global threshold of 58% computed using Otsu's method via MATLAB function "graythresh" still leads to inadequate performance of the global thresholding schema, suggesting that the histogram of intensities in the tumor volume is not truly bi-modal. The same conclusion can be drawn for local thresholding and for both clustering methods utilizing 2 clusters as they all operate under the assumption of bi-modal intensity distribution. However, the dynamic K-means clustering with 2 clusters (second only to SSRG/4D-PET) has a much higher Dice value than "static" K-means clustering with 2 clusters and a much lower misclassification rate due to all the extra pre-processing steps involved and information obtained from TACs. Because only those clusters containing voxels with highest mean values were classified as tumor, clustering with 4 clusters did not produce satisfactory results as expected. Best delineation results for both clustering algorithms were expected for clustering with 3 clusters, since the selected VOI contained three distinctly different tissue types. Our expectation was satisfied with regards to "static" K-means clustering. Much to our surprise, however, the dynamic clustering

misclassification rate improved only slightly over the 2 cluster case, while the value of Dice decreased by close to 6%. This may be attributed to the smoothing effect of various filtering procedures applied to the data in the pre-processing stage. As a result, the cluster with the highest mean value did not include all the tumor voxels. Also, the slope values for the liver tumor tissue were not as steep compared to the background in the selected VOI as for rectal cancers used in validation of the dynamic K-means clustering algorithm by *Janssen et al*.[32], mostly due to the small amount of skeletal muscle tissue included in the VOI. In comparison to other algorithms, SSRG/4D-PET has the highest Dice value and the lowest misclassification rate. In fact, no false positives were registered for the proposed algorithm. In contrast to the other methods compared, it requires all the voxels assigned to the tumor to be spatially connected. Denoising with PCA and the spatial connectivity requirement in combination with the stopping condition employed within the single seed region growing framework led to more accurate tumor volume delineation.

6.2 Experimental Validation

Experimental validation of the algorithm was performed on an Allegro-PET whole body PET scanner (Philips Medical Systems Inc.).

6.2.1 Experimental Dynamic 4D PET Data

Measurements were performed with a cylindrical phantom (long axis coincident with the reconstruction centre and orthogonal to the image plane; inside diameter 19.0 cm), which was filled with ^{11}C ($T_{1/2}$ = 20 min, A = 0.006 MBq/ml) to simulate background. A cylindrical target (inside diameter = 4.75 cm, inside length = 8.5 cm), filled with ^{18}F ($T_{1/2}$ = 110 min, A = 0.002 MBq/ml), was mounted in the centre of the phantom with long axis coincident with the reconstruction centre to simulate a tumor volume. A dynamic sequence of 26 frames (45 slices per frame) was acquired for 130 min according to a schedule of 26 time frames of 300 s duration each. The data was corrected for attenuation and

scatter, and then reconstructed using 3D-RAMLA (row-action maximum likelihood algorithm).

6.2.2 Experimental Study: Results and Discussion

This study aims to demonstrate the validity of the PCA procedure within the framework of the proposed algorithm (SSRG/4D-PET) and, to a lesser degree, to provide experimental validation under more realistic clinically relevant conditions with all the corrections applied. However, all the delineation methods compared in this experimental study are expected to produce satisfactory results due to the size, shape, and placement of the simulated target volume. A relatively large (total volume of 150.62 cc), cylindrically symmetric target placed in the centre of the phantom with long axis coincident with the reconstruction centre is reasonably simple to delineate. The true bi-modal nature of the intensity histogram in the masked volume proves advantageous for the histogram shape-based image thresholding techniques based on global threshold. The same is also true for clustering with the number of clusters restricted to 2.

The transverse, coronal, and sagittal projections through the center of the simulated tumor in the first time frame of the raw unprocessed image volume are presented in **Figure 6.8**. The noisy and not entirely uniform nature of the activity distribution in the raw data is quite evident (**Figure 6.8a**, **Figure 6.8b**, and **Figure 6.8c**) and serves to approximate (at least to first order) the combination of normal and tumor cells often found within a target volume.

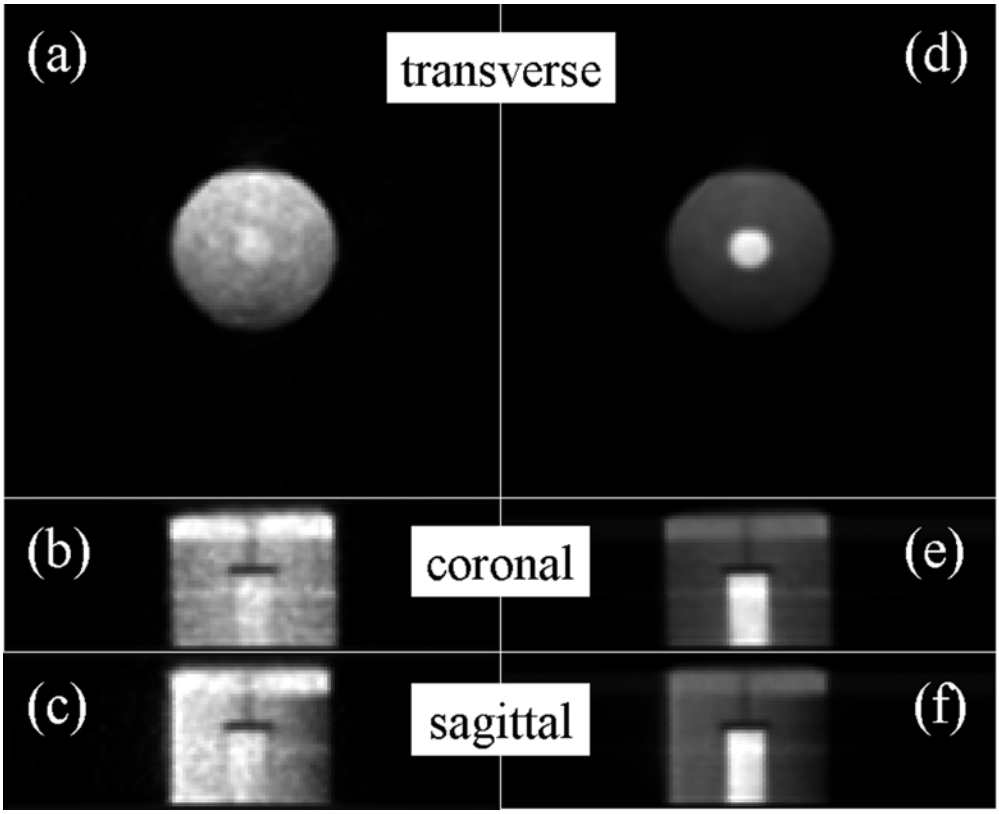

Figure 6.8: Comparison of raw unprocessed image and corresponding PC1 image: (a)-(c) transverse, coronal, and sagittal projections for the raw data; (d)-(f) transverse, coronal, and sagittal projections for the PCA processed data. (a), (d): slice 33, first time frame; (b), (c), (e), (f): projections through the center of simulated tumor.

The noise and artifact suppression provided by application of the first PCA channel (PC1 image) to the dynamic PET data in the first stage of the SSRG/4D-PET algorithm is clearly illustrated in **Figure 6.8**. The effectiveness of PCA applied to a dynamic PET scan for denoising and identification of structures is readily visible (**Figure 6.8d**, **Figure 6.8e**, and **Figure 6.8f**)[49, 82]. In this method the TAC of each voxel is individually compared to the TAC of the seed (through a scaling parameter). The noise present in TACs, if left untreated, can be detrimental to the algorithm's performance since the TAC model used in the parameterization does not account for noise. For this phantom, SSRG/4D-PET

applied to raw untreated data results in a 68.64% volume discrepancy between delineated and true target volumes, whereas for the PCA processed data the volume discrepancy is reduced to 12.34%.

Table 6.7: Phantom study – comparison of delineation results for six algorithms

Method	Dice coefficient	R_m (%)
SSRG/4D-PET	0.93	12.55
Dynamic K-means clustering (*Janssen et al.*[32]) with 2 clusters	0.92	15.40
K-means clustering with 2 clusters	0.92	14.96
Fixed threshold of maximum signal intensity (40%)	0.91	16.92
global automatic thresholding (Otsu[123])	0.92	15.49
adaptive (local) automatic thresholding (Otsu[123])	0.83	31.12

Table 6.7 presents segmentation results of six different approaches evaluated by Dice coefficient and misclassification rate (R_m). This evaluation was based on a comparison of the target volumes delineated using each of the algorithms to the manually delineated target volume. Manual target volume delineation was performed using the Region Of interest Visualization, Evaluation, and image Registration (ROVER) software package (ABX GmbH, Germany) based on known physical dimensions and placement of the cylindrical target. In order to compare the four static approaches to SSRG/4D-PET target volume delineation, the dynamic data set was averaged in the temporal direction prior to delineation. The target volume was delineated on the same masked volume used in the proposed algorithm (SSRG/4D-PET). From inspection of the data presented in

Table 6.7, SSRG/4D-PET delineation reveals a distinct advantage over all four alternate segmentation techniques. The "adaptive thresholding" method clearly underestimates the target volume and yields the highest misclassification rate compared to SSRG/4D-PET, followed closely by the "fixed threshold" method. It is noteworthy that "adaptive thresholding" performs quite poorly in comparison with the other methods. Adaptive thresholding was performed by applying a morphological top-hat operator to the data and the resulting image volume was processed using a global threshold obtained from Otsu's method. The combination of the smoothing effect of the procedure, size of the structuring element used (square, 8×8 pixels), and the non-negligible impact of the plastic walls of the cylindrical target on local threshold values may explain the degradation in performance of the "adaptive thresholding" method. Delineation with dynamic K-means clustering (2 clusters) results in a slightly higher misclassification rate (same Dice value) than does "static" K-means clustering with 2 clusters (**Table 6.7**). This can be attributed to the smoothing effect of various filtering procedures applied in the data pre-processing stage. In the case of both "K-means clustering" techniques and "global thresholding", the advantage of the SSRG/4D-PET method was less pronounced, but nonetheless evident (the highest Dice value and the lowest misclassification rate).

6.3 Clinical Evaluation

Clinical evaluation of the algorithm was performed on sets of images acquired with a Philips Gemini GS PET/CT scanner (Philips Medical Systems Inc.). To analyze the performance of the proposed algorithm (SSRG/4D-PET) using clinical data, the biological target volumes (BTVs) obtained are compared to the results of static delineation using automated SUV thresholding.

6.3.1 Clinical Patient Data

To evaluate the performance of the SSRG/4D-PET delineation algorithm in a clinical setting, data from four patients (patient ID: *1*, *2*, *3*, and *4*) diagnosed with prostate cancer were analyzed (**Table 6.8**). These patients were enrolled with

informed consent in a research protocol which had received local Research Ethics Committee approval. Both CT and dynamic PET images of the region of interest were acquired. Patients were intravenously injected with [C-11] Choline[m] immediately after beginning their PET scan in a "head-first supine" position with arms held above their head.

Table 6.8: Clinical evaluation – patient specific characteristics and comparison of delineation results for dynamic (SSRG/4D-PET; COF_{01} (section **6.3.3**)) and static (fixed threshold of maximum signal intensity, i.e. % of SUV_{max} thresholding; COF_{02} (section **6.3.3**)).

Patient ID	Age (y)	Weight (kg)	[C-11] Choline activity (MBq)	Dynamic target volume (cm^3)	Static target volume (cm^3)	Threshold for static delineation (% SUV_{max})	COF_{01}	COF_{02}
1	79	100.0	862	33.2	22.1	75	0.67	1.00
2	68	84.0	870	30.3	20.2	75	0.67	1.00
3	74	85.0	804	32.5	37.3	75	0.98	0.85
4	70	92.0	862	32.2	30.6	88	0.90	0.95

PET data was acquired in list mode for a total of 40 minutes. This data was re-binned into 20 sinograms, each containing 2 minutes of list mode data. Sinograms were reconstructed according to current local clinical protocol. Routine clinical image reconstruction using default (scanner manufacturer supplied) reconstruction tools was performed with a fast, fully 3D iterative algorithm (3D-RAMLA) using two iterations, a relaxation parameter of 0.006 and

[m] [C-11] Choline is currently the most promising tracer for imaging prostate cancer patients. M. Picchio, E. Giovannini, C. Crivellaro, et al., "Clinical evidence on PET/CT for radiation therapy planning in prostate cancer," Radiother.Oncol. 96 (3), 347-350 (2010).

a “blob” radius of 2.5 pixels. For each frame the reconstructed image size was 144×144 pixels in 45 slices with an isotropic voxel size of 4×4×4 mm^3. Scans were normalized to correct for variations in detector efficiency and distortion. Emission data was corrected for randoms (direct online subtraction), scatter (single scatter simulation – SSS), attenuation (using segmented CT data, CT-3DAC) and decay. Software used in the analysis of these reconstructed image data sets was developed in MATLAB (version 7.4.0.287, R2007a; The Mathworks, Natick, MA), on a PC.

6.3.2 Biological Target Volume Delineation

For each patient, a contour was derived from the dynamic scan data using the SSRG/4D-PET delineation algorithm. A second contour was obtained from a static data set (consisting of the average of the last 15 dynamic time frames) using automated SUV thresholding. For this purpose, the software package "Region Of interest Visualization, Evaluation, and image Registration" (ROVER, ABX GmbH, Germany) was used. The threshold used in static delineation was based on a percentage of the maximum SUV value (SUV_{max}) within a spherical mask with diameter of 17 pixels (6.8 cm) centered on the prostate. This $\%SUV_{max}$ was set by an experienced staff physician to 88% for patient *4* and to 75% for the other patients (*1*, *2*, and *3*).

In static PET analysis, only volumes corresponding to the last 15 frames (last 30 minutes of acquisition) were averaged and this time averaged data was used for biological target volume (BTV) delineation. For the analysis of both static and dynamic PET data, the mask manually chosen by the user was of a fixed form: a sphere for static PET analysis and cuboid for dynamic PET analysis. **Figure 6.9** and **Figure 6.10** depict transverse, coronal, and sagittal projections taken through the seed point (location indicated by cross hairs) used for delineation with SSRG/4D-PET and showing a cross section of the spherical VOI used in static delineation (SUV_{max} thresholding). For SSRG/4D-PET a fully enclosing VOI cuboid was used.

3D binary datasets (having tumor voxels labeled with a value 1 and zero everywhere else) resulting from both dynamic and static PET data analyses were converted to contours on each slice using the MATLAB function "contour" and were used for visualization purposes. Volumes defined by these binary masks were used for algorithm performance evaluation.

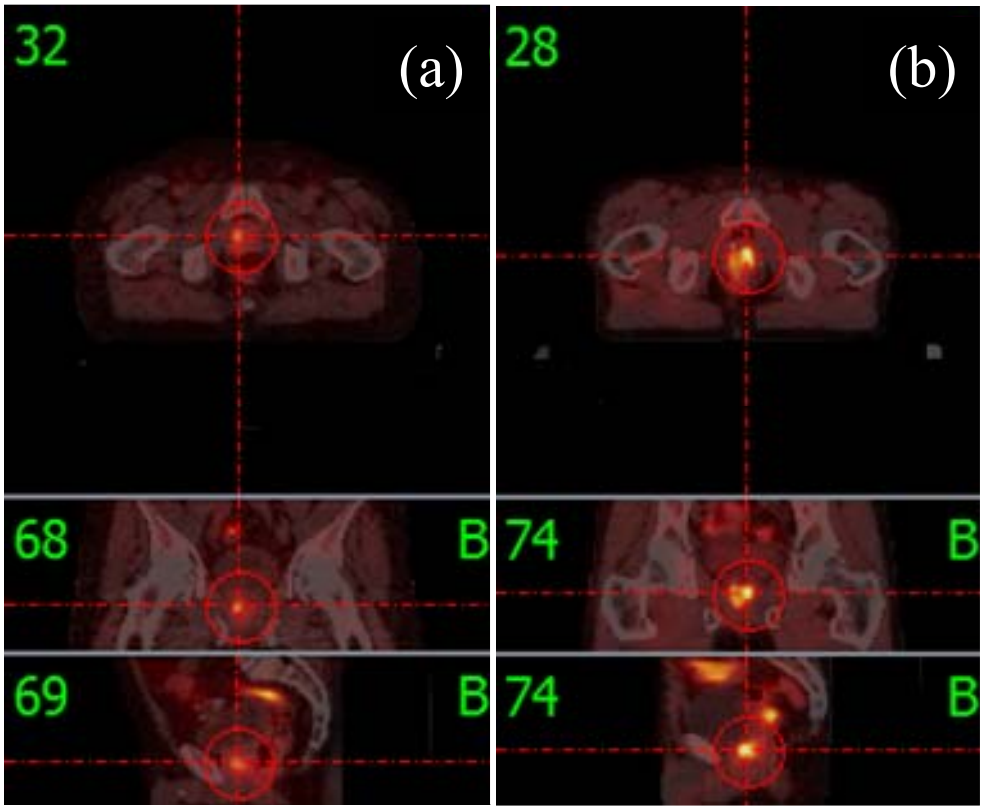

Figure 6.9: (a) – patient *1* and (b) – patient *2*. Fused PET/CT images (transverse, coronal, and sagittal slices) showing the spherical VOI used in static delineation (SUV_{max} thresholding) and the coordinate of the seed used for dynamic (SSRG/4D-PET) delineation. For SSRG/4D-PET, a cuboid fully enclosing the VOIs shown here was used.

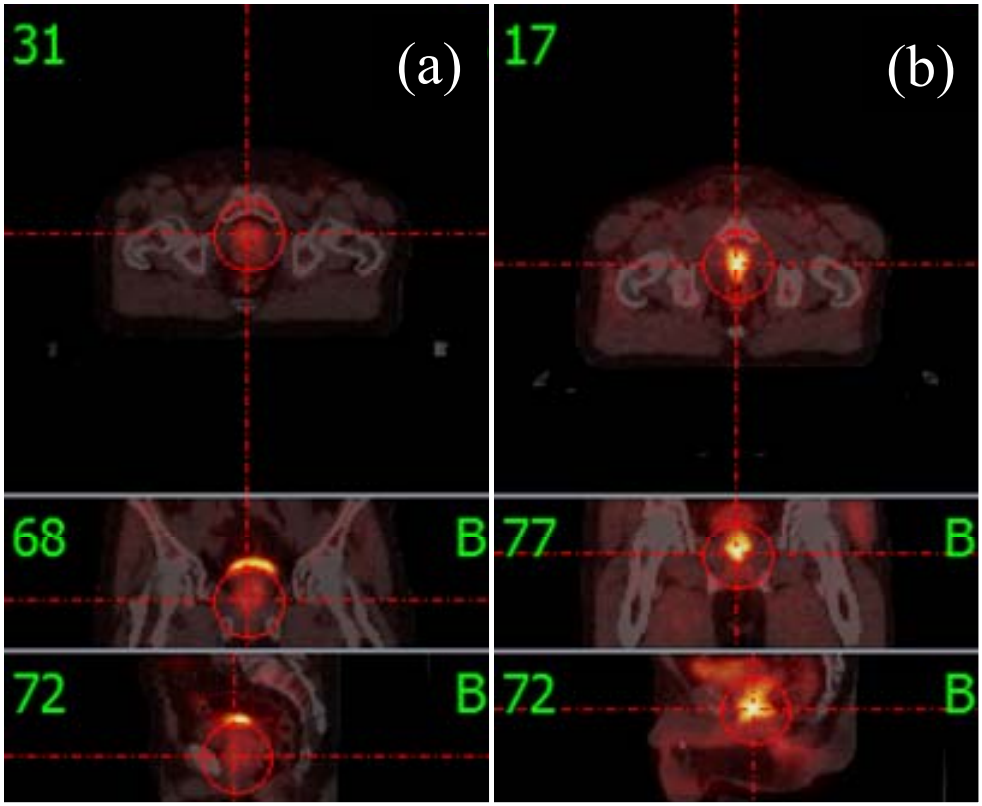

Figure 6.10: (a) – patient *3* and (b) – patient *4*. Fused PET/CT images (transverse, coronal, and sagittal slices) showing the spherical VOI used in static delineation (SUV_{max} thresholding) and the coordinate of the seed used for dynamic (SSRG/4D-PET) delineation. For SSRG/4D-PET, a cuboid fully enclosing the VOIs shown here was used.

6.3.3 Contour Comparison Methods

To investigate the differences between the contours (C_1 and C_2) resulting from these two delineation approaches, the method of analysis used by *Janssen et al*.[32] was followed. Three regions were defined as depicted on **Figure 6.11**.

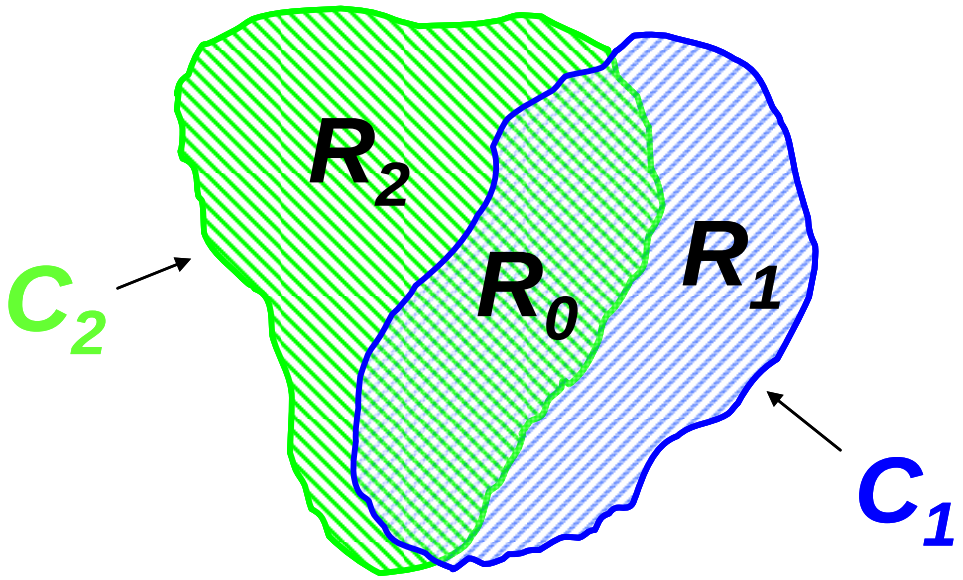

Figure 6.11: R_1 – upwards diagonal (from left to right), R_2 – downwards diagonal (from left to right) and R_0 – crosshatched region (intersection of contours C_1 and C_2). Schematic diagram depicting three regions (R_0, R_1, and R_2) used in calculation of contour overlap fractions (COF_{01} and COF_{02}) for two contours (C_1 and C_2).

Region R_1 represents that part of contour C_1 not overlapping with contour C_2 and Region R_2 represents that part of contour C_2 not overlapping with contour C_1, whereas region R_0 is defined as the intersection of contours C_1 and C_2. Two contour overlap fractions (COFs) were calculated using the number of voxels in the volumes (summed over all slices in which these contours occurred) corresponding to these three regions (equations **6.13** and **6.14**). The overlap fraction with respect to the first contour is defined as:

$$COF_{01} = \frac{N_{region\ R_0}}{N_{region\ R_0} + N_{region\ R_1}} \qquad \textbf{6.13}$$

where $N_{region\ R_0}$ and $N_{region\ R_1}$ are the number of voxels in regions R_0 and R_1. In turn, the overlap fraction with respect to the second contour is calculated as:

$$COF_{02} = \frac{N_{region\ R_0}}{N_{region\ R_0} + N_{region\ R_2}} \quad \textbf{6.14}$$

where $N_{region\ R_0}$ and $N_{region\ R_2}$ are the number of voxels in regions R_0 and R_2. The differences between the contours are assessed based on these overlap fractions.

To fully analyze algorithm performance, detailed analysis of the shapes of the TACs within non-overlapping regions as well as within the intersection of the two contours was performed. It is reasonable to assume that the shape of voxel TACs within the intersection region (R_0) of the two contours is characteristic of malignant tissue. Based on this assumption, and using the TACs of the intersection region as indicative of tumor, the voxels within non-overlapping regions (R_1 and R_2) are classified as cancerous or normal tissue depending on the shape of their corresponding TACs.

6.3.4 Statistical Analysis

Significant differences between biological target volume delineation by dynamic (SSRG/4D-PET) and static (fixed threshold as percentage of maximum SUV) approaches were expected. For the evaluated patient population, dynamic and static target volumes obtained are expressed in terms of mean ± standard deviation and range. To test the statistical significance of observed differences, a Wilcoxon signed-ranks test for the median difference between the target volumes resulting from the two tumor delineation methods was performed. The MATLAB function "signrank" (Statistics Toolbox; Version 5.0.1) performs a paired, two-sided test of the hypothesis that the difference between the matched items comes from a distribution whose median is zero. In this test, the level of significance was selected at 0.05 ($\alpha = 0.05$), meaning that the observed differences are considered statistically significant if the obtained p value is less than 0.05.

6.3.5 Tissue-Specific Uptake Kinetics

An important assumption, upon which the delineation algorithm (SSRG/4D-PET) is predicated, is the existence of significant differences between the time-activity curves of malignant tissue as compared to normal tissue. To confirm the validity

of this assumption we have examined TAC patterns for different tissue types identified within the four patients available to us (**Figure 6.12**, **Figure 6.13**, **Figure 6.14**, and **Figure 6.15**).

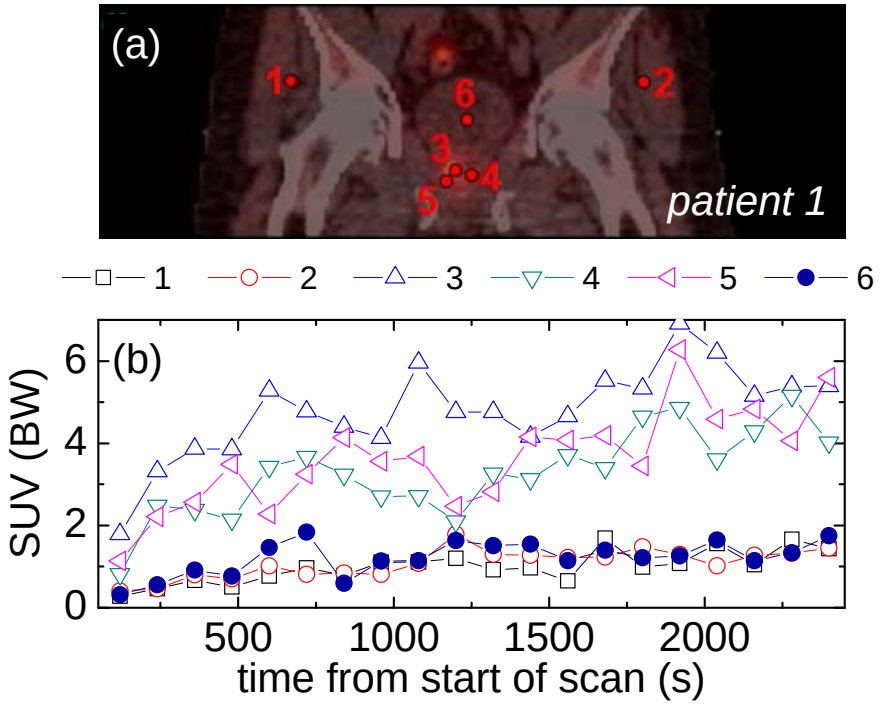

Figure 6.12: Fused PET/CT image at the last time frame of slice 20 (coronal projection) is presented in (a) for patient *1*; from which six points were selected manually within different tissue types: normal tissue (1, 2), tumor tissue (3, 4), one point at the boundary of the tumor (5), and one point within the bladder (6). The TACs of the selected points are shown in (b).

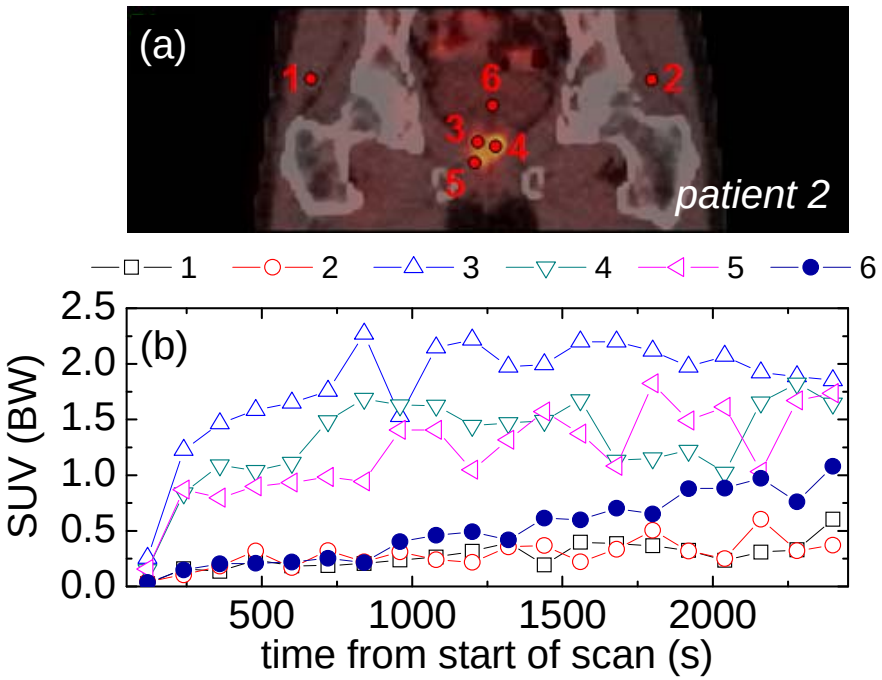

Figure 6.13: Fused PET/CT image at the last time frame of slice 20 (coronal projection) is presented in (a) for patient *2*; from which six points were selected manually within different tissue types: normal tissue (1, 2), tumor tissue (3, 4), one point at the boundary of the tumor (5), and one point within the bladder (6). The TACs of the selected points are shown in (b).

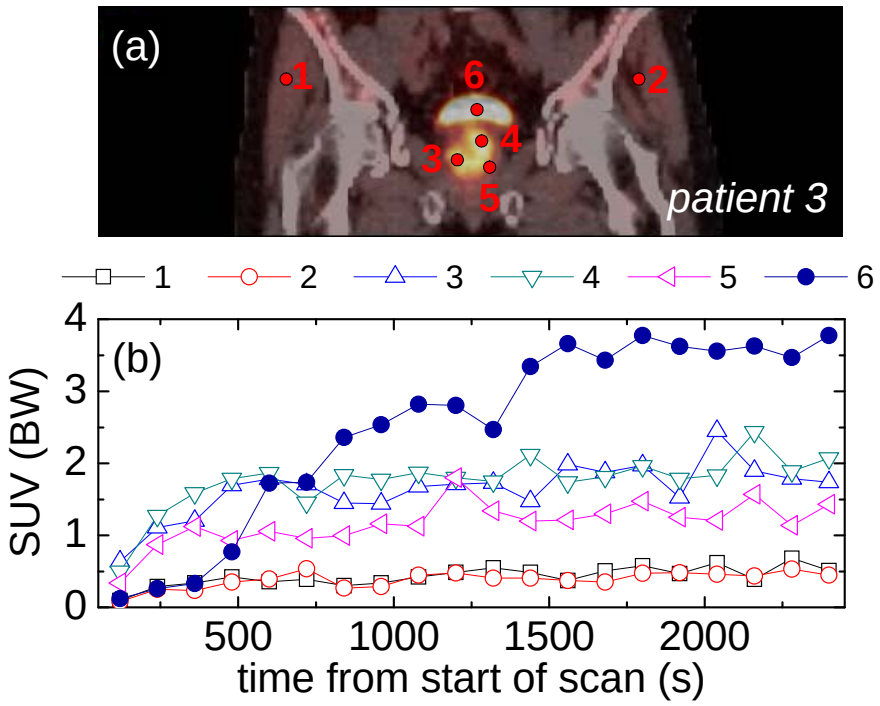

Figure 6.14: Fused PET/CT image at the last time frame of slice 20 (coronal projection) is presented in (a) for patient *3*; from which six points were selected manually within different tissue types: normal tissue (1, 2), tumor tissue (3, 4), one point at the boundary of the tumor (5), and one point within the bladder (6). The TACs of the selected points are shown in (b).

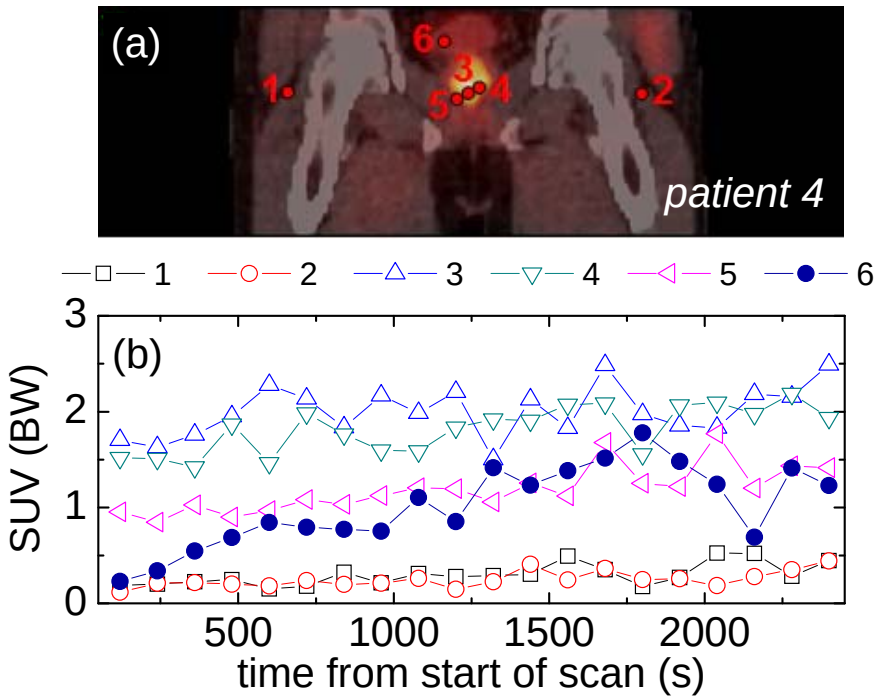

Figure 6.15: Fused PET/CT image at the last time frame of slice 20 (coronal projection) is presented in (a) for patient *4*; from which six points were selected manually within different tissue types: normal tissue (1, 2), tumor tissue (3, 4), one point at the boundary of the tumor (5), and one point within the bladder (6). The TACs of the selected points are shown in (b).

Points 1 and 2, located in uninvolved tissues of all four patients reveal a relatively small increase in accumulated activity as a function of time. In contrast, voxels lying within the prostate for patients *1*, *2*, and *3* (**Figure 6.12**, **Figure 6.13**, and **Figure 6.14**) exhibit an initial rapid increase in activity level (well above those of points 1 and 2) over the first 3 to 4 time frames (first 6 to 8 minutes of their scans) followed by a plateau of relatively constant uptake. The uptake pattern within the prostate indicates that the scan of patient *4* (**Figure 6.15**) did not begin promptly but rather several minutes post injection thus capturing the plateau phase only. The time activity curves for the three points within the prostate also reveal

significant amounts of variation as a function of time for all four patients. This is in all probability due to patient motion during these scans in combination with the presence of TAC heterogeneity within the prostate. A varied response is also observed within the bladder for these four patients. A large rise in SUV is observed for patient *3*, a more moderate rise for patient *4*, a small rise for patient *2*, and for patient *1* no discernable rise different from that of points 1 and 2. These variations in urinary accumulation of [C-11] Choline are suspected to be a result of incomplete tubular re-absorption of intact tracer and/or enhanced excretion of labeled oxidative metabolites[175]. Distinct differences between the shapes of the TACs, evident from the data presented in **Figure 6.12**, **Figure 6.13**, **Figure 6.14**, and **Figure 6.15**, suggest tissue specific Choline uptake kinetics.

6.3.6 Contour Comparison Results

The average biological target volume obtained using our algorithm (SSRG/4D-PET) was *32.0±1.2 cm*^{3} (range, *30.3–33.2 cm*^{3}), whereas for static delineation (standardized uptake value (SUV) thresholding based on percentage of SUV_{max}) the average resulting volume was *27.6±7.9 cm*^{3} (range, *20.2–37.3 cm*^{3}). An overview of the patient characteristics and comparison of results for dynamic and static delineation is presented in **Table 6.8**.

Contours resulting from dynamic (SSRG/4D-PET) and static target volume delineation are overlaid on their corresponding PET images (representative transverse slices) and are presented in **Figure 6.16**, **Figure 6.17**, **Figure 6.18**, and **Figure 6.19** for all four patients examined in this study. Images in these figures are cropped to enhance visibility of the differences between contours.

Analysis of the differences between the contours resulting from dynamic delineation using SSRG/4D-PET and contours generated from static SUV thresholding revealed contours obtained with dynamic (SSRG/4D-PET) PET analysis to be on average significantly larger (*20.5±34.1%,* range: *-12.9 to 50.2%*, $p < 0.0045$) than their static counterparts. *80.3±16.1%* (range, *66.6–98.0%*) of the dynamic (SSRG/4D-PET) target volume was located within the contour resulting from static SUV analysis. In contrast, *95.0±6.9%* (range, *85.4–100.0%*) of the

static SUV contour was located within the contour resulting from dynamic analysis (**Table 6.8**).

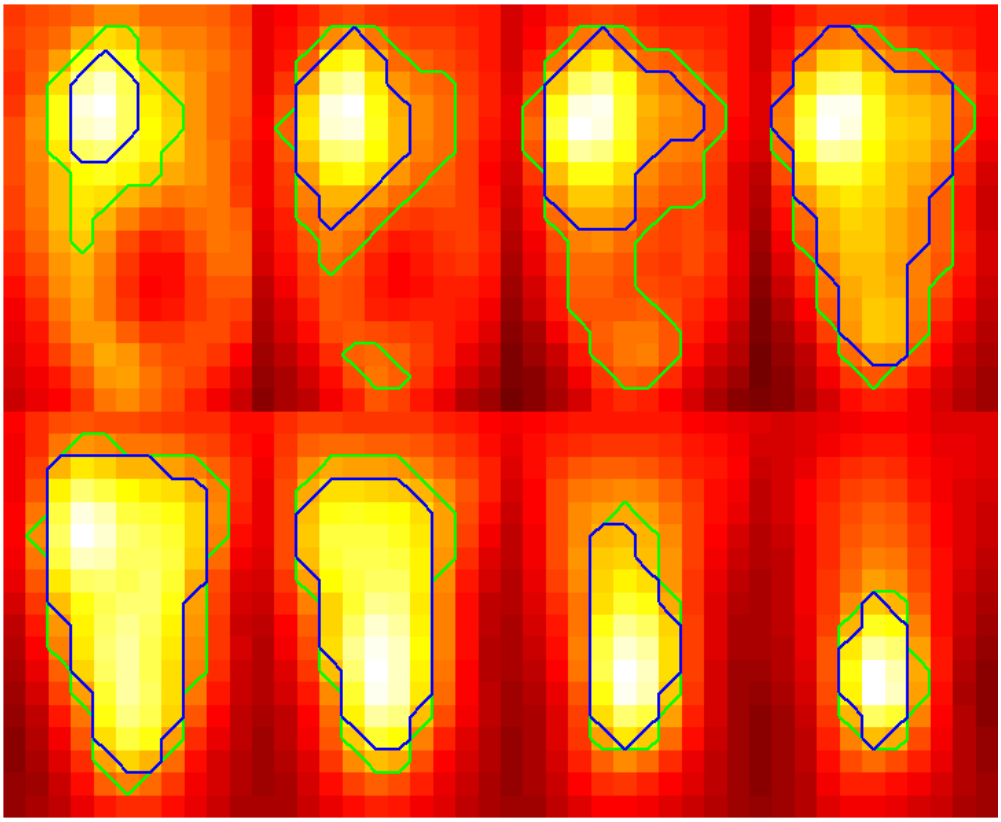

Figure 6.16: Patient *1*, representative slices at the last time frame. In blue, the target contour resulting from standardized uptake value thresholding (75% of SUV_{max}) of static PET data; in green, the target contour resulting from dynamic PET delineation (SSRG/4D-PET).

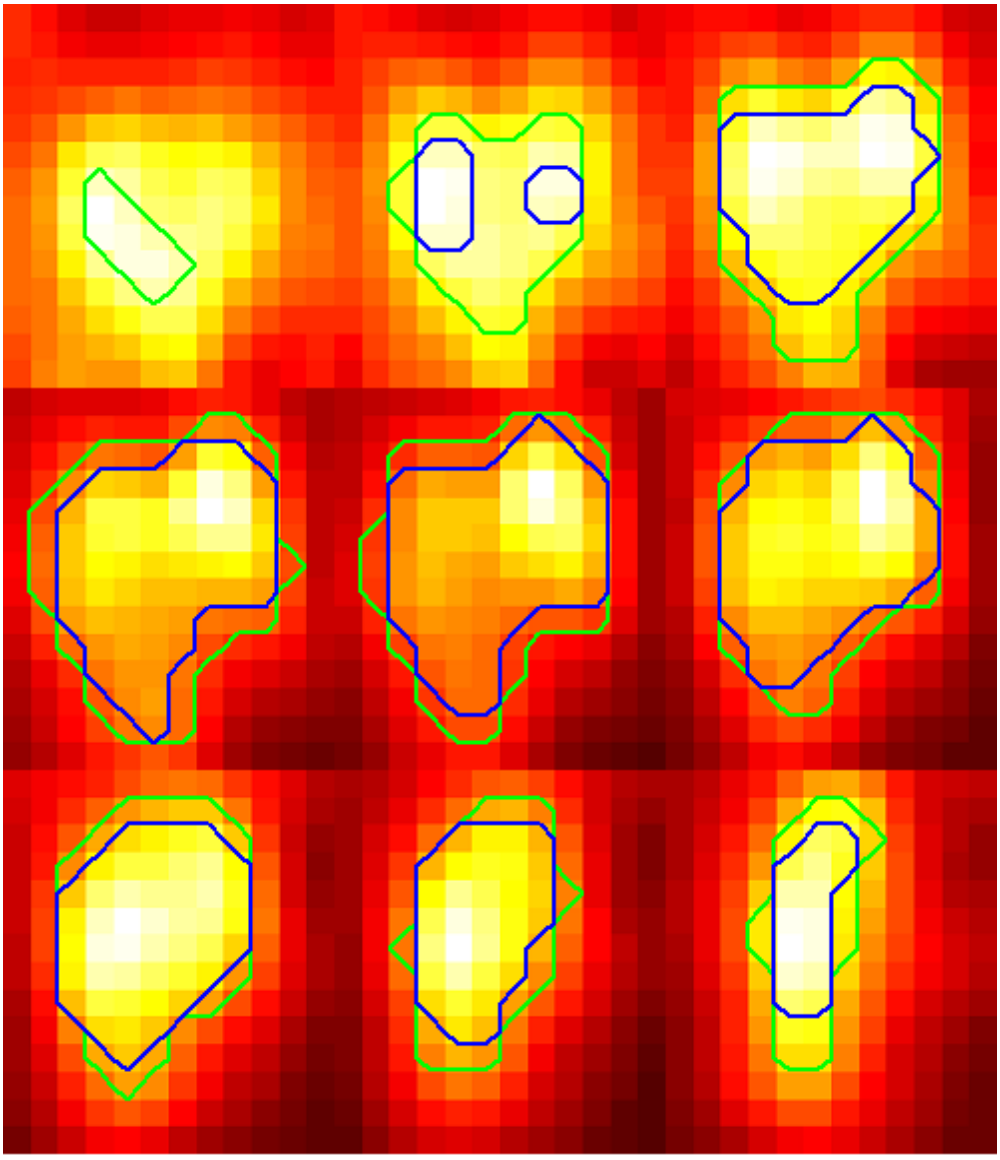

Figure 6.17: Patient *2*, representative slices at the last time frame. In blue, the target contour resulting from standardized uptake value thresholding (75% of SUV_{max}) of static PET data; in green, the target contour resulting from dynamic PET delineation (SSRG/4D-PET).

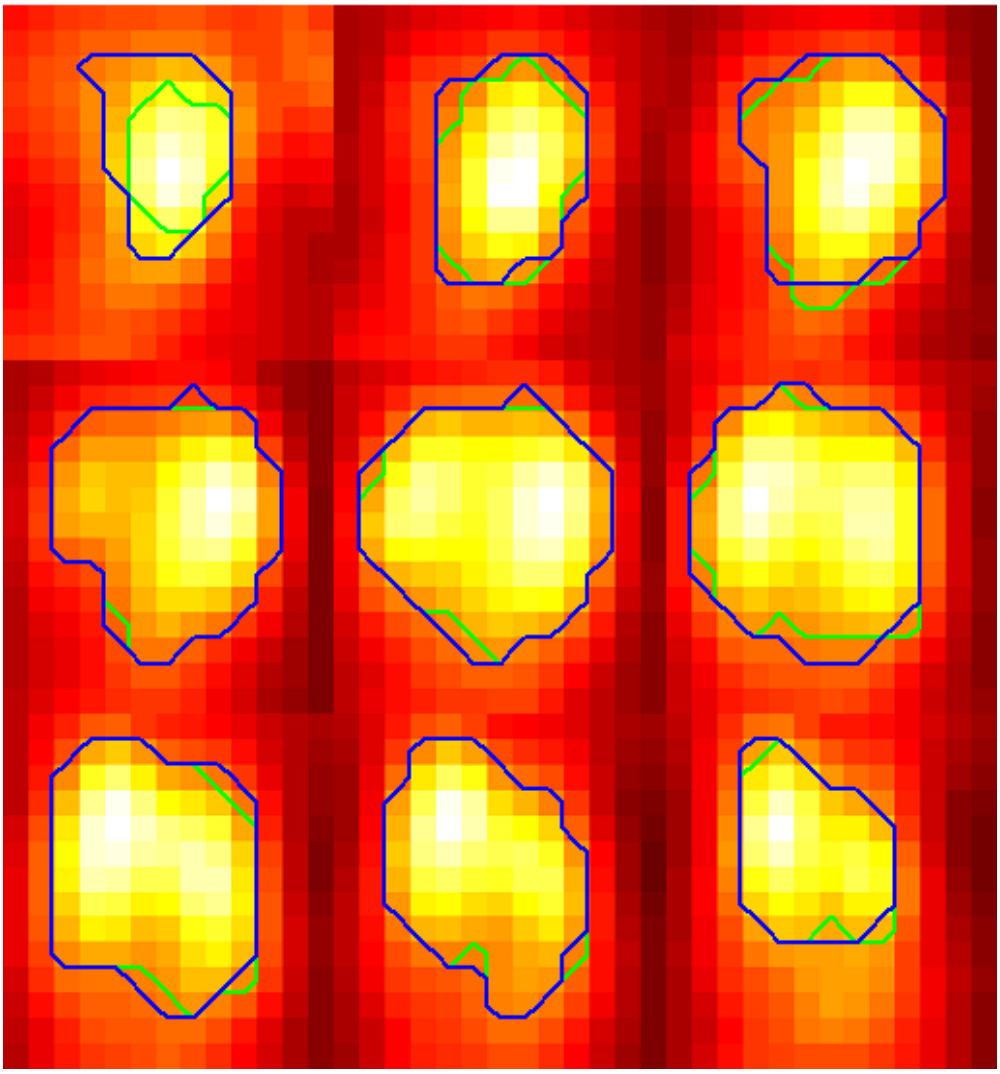

Figure 6.18: Patient *3*, representative slices at the last time frame. In blue, the target contour resulting from standardized uptake value thresholding (75% of SUV_{max}) of static PET data; in green, the target contour resulting from dynamic PET delineation (SSRG/4D-PET).

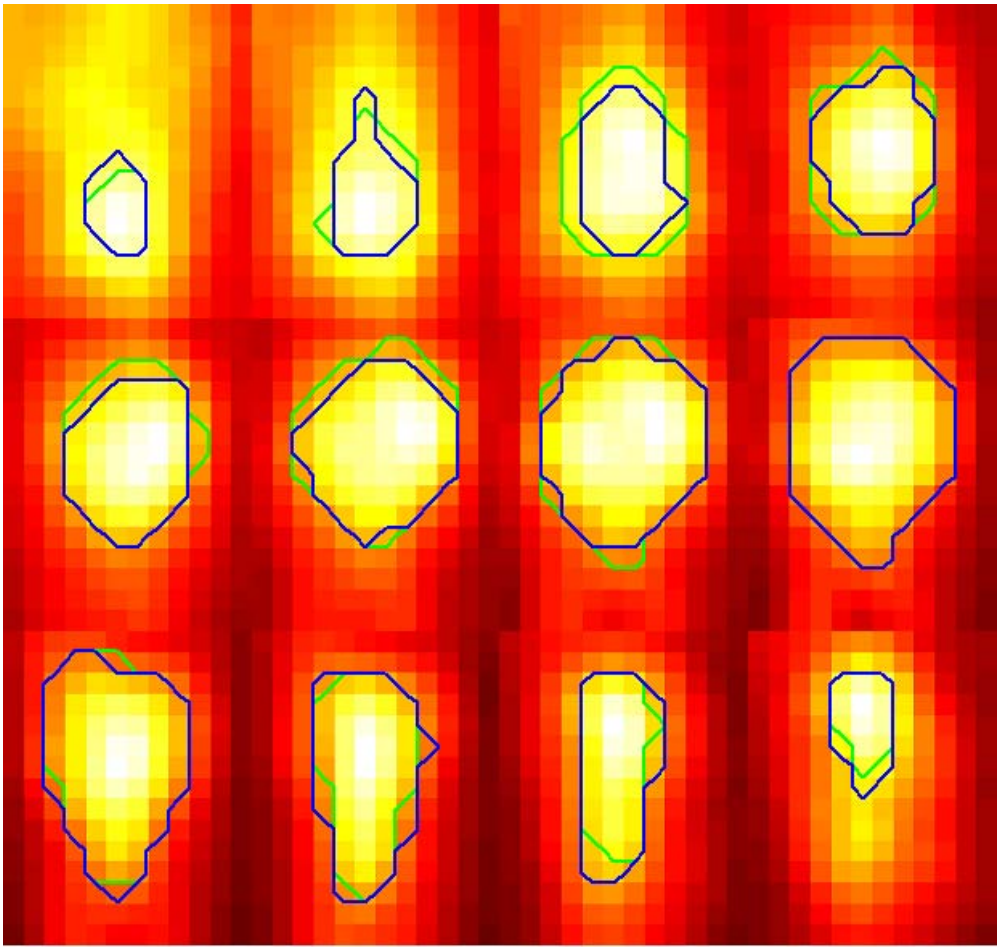

Figure 6.19: Patient *4*, representative slices at the last time frame. In blue, the target contour resulting from standardized uptake value thresholding (88% of SUV_{max}) of static PET data; in green, the target contour resulting from dynamic PET delineation (SSRG/4D-PET).

A comparison of the time activity curves (TACs) within the contours which result from dynamic (SSRG/4D-PET) and static (standardized uptake value (SUV) thresholding based on a percentage of the maximum SUV) target volume delineation for patients *1* and *2* is shown in **Figure 6.20**. For both of these patients, the static contours (C_1) are fully circumscribed by the dynamic contours (C_2) so that R_0 (intersection region) is equal to C_1. Thus, for these geometries, it is not surprising that the TACs within the static contours exhibit the same shape as those characterizing the dynamic contours. These plots (**Figure 6.20**) are

extremely busy as they contain the TACs of all voxels within these contours. To aid in interpretation, the mean TAC over all TACs in each contour are presented in **Figure 6.21**. The error bars on each mean TAC are derived from the standard deviation over the range of TACs from each region. The similarity between the shapes of the TACs in each contour clearly indicates that static delineation failed to recognize the full extent of the tumor. Treatment plans for these two patients which were based solely on the results of static delineation would result in failure to target the entire extent of the tumor.

The time activity curves within the static and dynamic contours for patients *3* and *4* are presented in **Figure 6.22**. For these two patients the static contours (C_1) were not fully circumscribed by their dynamic counterparts (C_2). TACs for the intersection regions (R_0) are presented in **Figure 6.22a** and **Figure 6.22d** (patients *3* and *4* respectively). TACs for the R_2 regions (within the dynamic contours but not intersecting with the static contours) are presented in **Figure 6.22b** and **Figure 6.22e**. TACs for the R_1 regions (within the static contours but not intersecting with the dynamic contours) are presented in **Figure 6.22c** and **Figure 6.22f**. For clarity of presentation the mean TACs over these regions (along with their standard deviations) are presented in **Figure 6.23**. The similarity between the TACs of regions R_0 and R_2 are evident. Also evident are the dissimilarities between the TACs of region R_1 and those of R_0. If, as hypothesized, the TACs of region R_0 are indicative of malignant tissue, for these two patients static delineation has failed to encompass the full extent of the tumor while at the same time including non-malignant tissues within the target volume. Therapeutic failure and the risk of elevated normal tissue complications are a possibility for treatment planning based solely on these static contours.

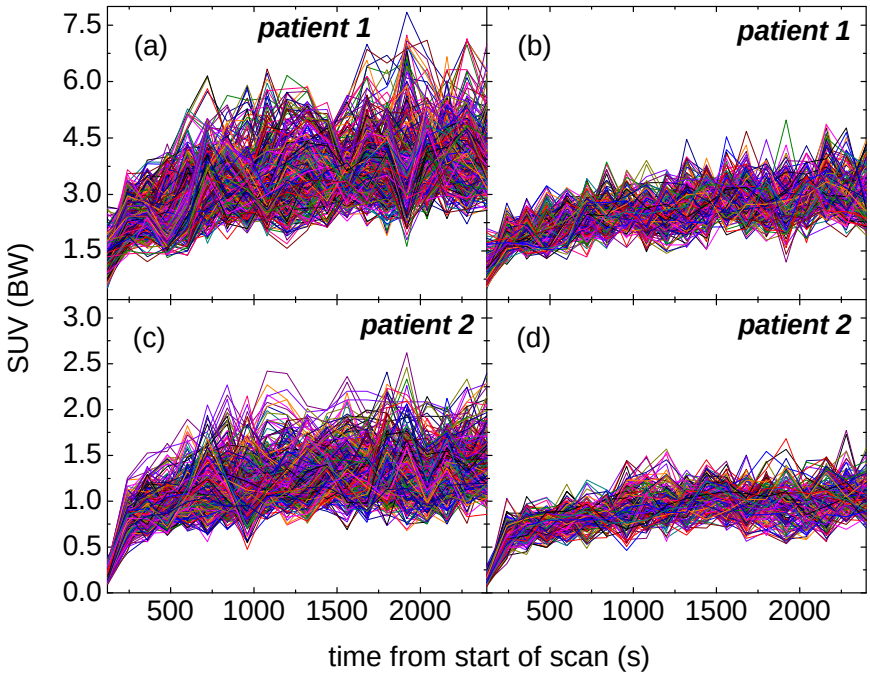

Figure 6.20: Comparison of time activity curves (TACs) within dynamic (SSRG/4D-PET) and static (thresholding, based on percentage of SUV_{max}) contours; (a), (b) - patient *1* and (c), (d) - patient *2*. (a), (c) - TACs of the voxels of all slices present within the intersection region of the two compared contours for respective patients. (b), (d) - TACs of the voxels of all slices present inside the dynamic contour but outside the static PET tumor contour.

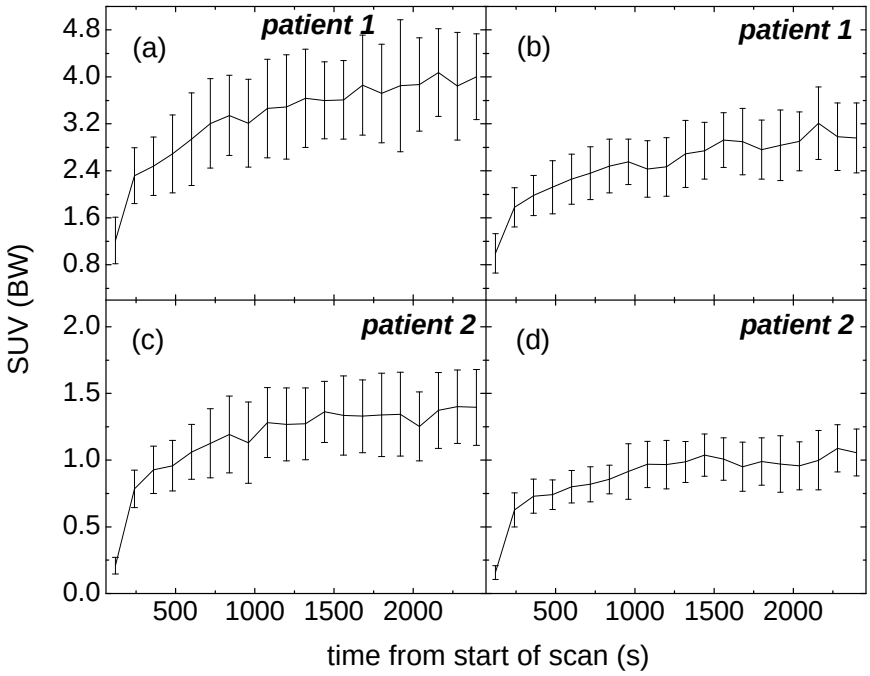

Figure 6.21: Comparison of average time activity curves (TACs) (with standard deviation plotted as error bar) within dynamic (SSRG/4D-PET) and static (thresholding, based on percentage of SUV_{max}) contours; (a), (b) - patient *1* and (c), (d) - patient *2*. (a), (c) - TACs of the voxels of all slices present within the intersection region of the two compared contours for respective patients. (b), (d) - TACs of the voxels of all slices present inside the dynamic contour but outside the static PET tumor contour.

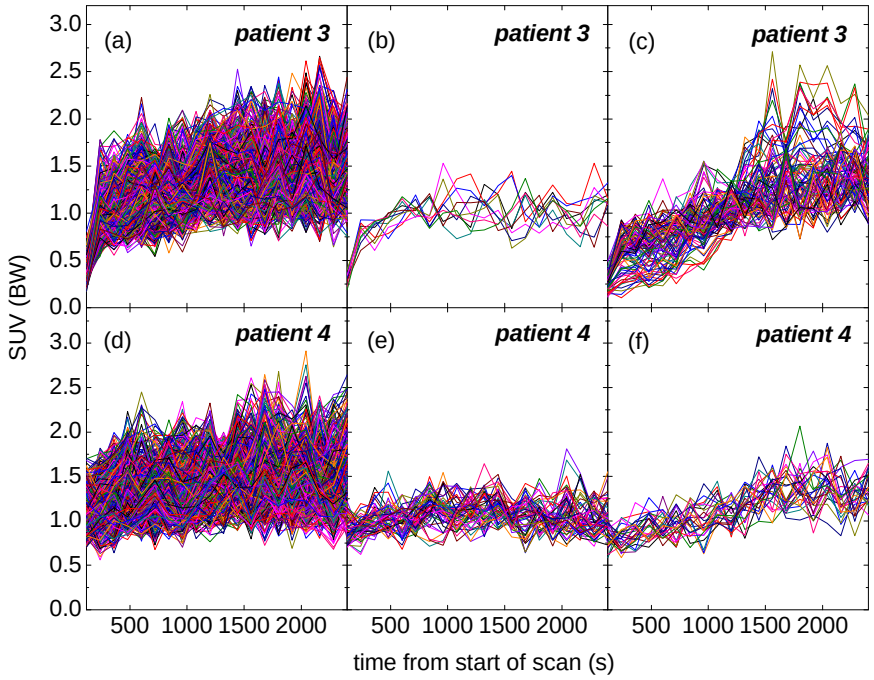

Figure 6.22: Comparison of time activity curves (TACs) within dynamic (SSRG/4D-PET) and static (thresholding, based on percentage of SUV_{max}) contours; (a), (b), and (c) - patient *3* and (d), (e), (f) - patient *4*. (a), (d) - TACs of the voxels of all slices present within the intersection region of the two compared contours for respective patients. (b), (e) - TACs of the voxels of all slices present inside the dynamic contour but outside the static PET tumor contour. (c), (f) - TACs of the voxels of all slices present inside the static contour but outside the dynamic PET contour.

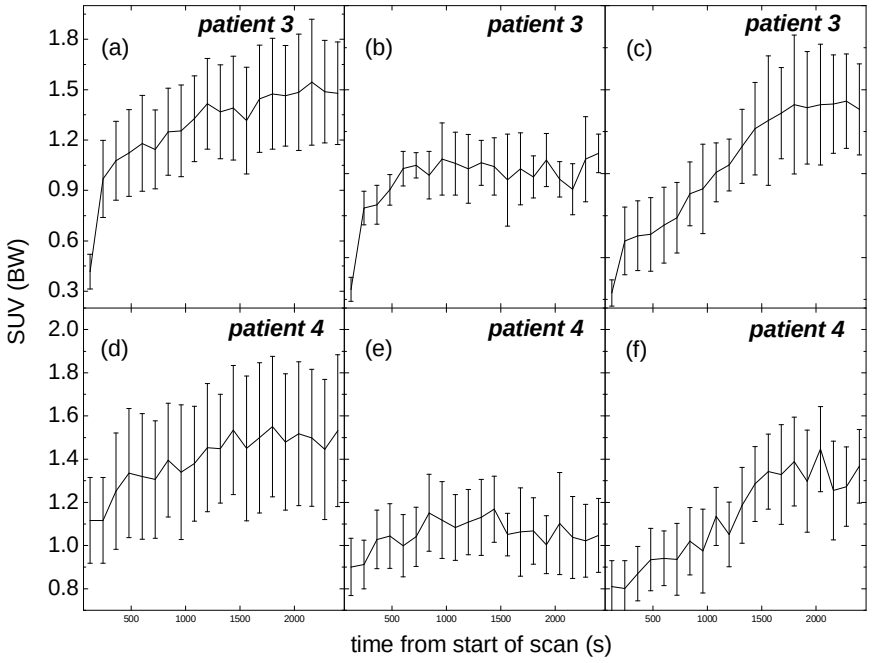

Figure 6.23: Comparison of average time activity curves (TACs) (with standard deviation plotted as error bar) within dynamic (SSRG/4D-PET) and static (thresholding, based on percentage of SUV_{max}) contours; (a), (b), and (c) - patient *3* and (d), (e), (f) - patient *4*. (a), (d) - TACs of the voxels of all slices present within the intersection region of the two compared contours for respective patients. (b), (e) - TACs of the voxels of all slices present inside the dynamic contour but outside the static PET tumor contour. (c), (f) - TACs of the voxels of all slices present inside the static contour but outside the dynamic PET contour.

6.3.7 Discussion

The preceding clinical evaluation is predicated upon the assumption that the region of intersection of static and dynamic (SSRG/4D-PET) contours serves to define a reasonable volume within which to define time activity curves which

accurately characterize cancerous tissue. Lacking any a-priori knowledge of the true extent of the tumor, as is the case with almost every in-vivo clinical study, this seems a reasonable assumption to make as static contours serve as the current standard for PET defined target volumes. Proceeding thusly, it has been shown that static delineation, at least for the four patients available for this limited endeavor, apparently fails to determine the full extent of tumor volumes as defined by time activity analysis. If true, these results indicate that target volumes defined solely upon static PET delineation run the risk of increasing the probability of compromising therapeutic benefit. Furthermore, in the case of patients *3* and *4*, static delineation also resulted in the inclusion of uninvolved tissues within target volumes. The inclusion of uninvolved tissues within target volumes may increase the risk of normal tissue complications which further detracts from the therapeutic benefit intended.

6.4 SSRG/4D-PET: Conclusions

The ability to accurately delineate target volumes which include the entire extent of malignant tissue while excluding as much normal tissue as possible is highly desirable if the metabolic information contained in PET images is to be used to its maximum benefit in radiation treatment planning.

A novel partially-supervised tumor volume delineation technique for dynamic PET imaging which allows improved classification of a patient's anatomy into cancerous and normal tissue in the presence of homogeneous and heterogeneous uptake was developed and evaluated using both synthetic and real clinical patient data. Although the main purpose of the technique is to improve delineation of the geometric extent of PET defined tumor volumes for use in PET-guided radiation therapy treatment planning, other applications in clinical oncology such as the assessment of response to treatment are also possible.

7. Summary and Future Work

In this final chapter we summarize the thesis (section **7.1**) and suggest future research directions (section **7.2**) which have the potential to enhance the value of the contributions made by this research effort.

7.1 Summary

4D-PET is garnering great interest in the radiotherapy community for its enhanced ability to better define target volumes. Significant foundational work, such as is presented in this thesis, is still required, however, to transform this potential into clinical reality. To this end we present a novel four dimensional (4D) tumor volume delineation algorithm (SSRG/4D-PET) that incorporates information available in dynamic PET imaging through time activity curves (TACs). This algorithm utilizes temporal and spatial data simultaneously via parametric images extracted directly from TACs for the purpose of tumor volume delineation.

The accuracy of tumor volume delineation is directly dependent on image quality. One of the confounding factors in the interpretation of PET data is image noise. Post-processing may be employed to reduce the level of noise present in clinical PET images. Effective image noise reduction is greatly dependent on an accurate knowledge of the parameters which characterize this noise. For this reason we first analyzed the statistical properties of noise in PET images reconstructed with Filtered-Backprojection (FBP) and Row-Action Maximum Likelihood Algorithm (RAMLA), after all clinical correction and image reconstruction procedures had been applied. From this analysis it was revealed that the noise in PET images created with RAMLA reconstruction is very well characterized by Gamma distribution followed closely by Normal distribution, while FBP produces comparable conformity with both Normal and Gamma statistics. We have also shown that the noise NSD (Normalized Standard Deviation = STD/Mean) is modeled equally well by the Negative Binomial, Normal, Log-Normal and Gamma distributions for RAMLA and by the Negative Binomial, Normal, and Gamma distributions for FBP reconstructions. These

results have been published in the Journal of Digital Imaging (*Teymurazyan et al*. 2012)[34].While radioactive decay is well-modeled as a Poisson process, the net result after all corrections and image reconstruction techniques have been applied is decidedly non-Poisson. A large number of noise reduction techniques are predicated on additive noise models and the incorrect treatment of image noise can be detrimental to optimal algorithm performance. Noise reduction algorithms specifically designed for Poisson noise are expected to produce inferior results when applied to clinical PET images.

The next step in the algorithm development was the integration of an effective post-processing technique aimed at noise reduction. A comparison between PET image denoising with different techniques (PCA, BLS-GSM, and SHINE) reveals PCA to be the most appropriate method for post-processing of dynamic PET data aimed at noise reduction prior to tumor volume delineation. PCA applied to dynamic PET image data is used to generate images with higher quality and improved SNR without relying on modeling assumptions. Within the framework of the proposed algorithm, retaining only the first principal component (PC1) allows obtaining data sets with a much lower levels of noise while enhancing contrast, thereby optimizing signal-to-noise ratio (SNR). The proposed algorithm (SSRG/4D-PET) utilizes temporal and spatial data simultaneously via parametric images extracted directly from TACs for the purpose of tumor volume delineation. The denoising step prior to parameterization is necessary because otherwise local short term fluctuations in TACs may significantly bias the results. For example, in the experimental validation study, SSRG/4D-PET applied to raw untreated data results in a 68.64% volume discrepancy between delineated and true target volumes, whereas with PCA processed data the volume discrepancy is reduced to 12.34%.

The ability to accurately delineate a target volume completely covering malignant tissue while including minimal amounts of normal tissue is necessary if the metabolic information contained in PET images is to be used to its maximum benefit in radiation treatment planning. A novel partially-supervised tumor volume delineation technique for dynamic PET imaging that allows improved

classification of the imaged volume into cancerous and normal tissue in the presence of homogeneous and heterogeneous uptake was developed and evaluated using synthetic and real clinical patient data. The performance of the proposed algorithm (SSRG/4D-PET) was evaluated using both semi-empirical digital phantoms (homogeneous and heterogeneous) and a clinically-relevant phantom imaged on an Allegro-PET whole body PET scanner. Both digital simulations and the physical phantom study suggest that accurate tumor volume delineation can be achieved with SSRG/4D-PET. Also, these simulations indicate that accurate delineation results with SSRG/4D-PET and short imaging intervals (25-30 min in total) may be realized. These shorter imaging times could help with balancing patient comfort and financial cost. For a more rigorous evaluation of the algorithm performance using clinical data, biological target volumes (BTVs) obtained from dynamic PET analysis (SSRG/4D-PET) were compared to the results of the current clinical practice of static delineation using automated SUV thresholding. Volumes of the contours resulting from dynamic delineation using SSRG/4D-PET were found to be on average significantly larger than contours generated from static SUV thresholding. On the other hand, in some of the evaluated patients, not all voxels showing TACs with shapes characteristic of what was observed in the region suspected to be malignant were present within the contour generated from static SUV analysis, introducing the possibility of geographic miss of malignant tissue. These voxels were, however, included in the contour resulting from dynamic delineation using SSRG/4D-PET.

In summary, SSRG/4D-PET exhibits promising potential to replace manual tumor delineation in clinical settings. PET-derived tumor volumes can be used to augment segmentation in CT (or MRI) images. Image fusion techniques may be used to compare anatomical information available through CT (or MRI) to the metabolically active tumor volumes delineated with SSRG/4D-PET, thus improving consistency and accuracy of the treatment planning process. With the advent of combined PET/CT systems, involvement of PET in the RT treatment planning process has been greatly simplified. The incorporation of 4-D PET into PET/CT offers great potential for improved tumor delineation. Also, it would

naturally lend itself to scans that are already acquired dynamically. Furthermore, even though the main purpose of the technique is to improve delineation of the geometric extent of PET defined tumor volumes for use in PET-guided radiation therapy treatment planning, other applications in clinical oncology such as the assessment of response to treatment are also possible.

7.2 Future Work

In this thesis, a new approach to the delineation of radiation therapy target volumes using dynamic PET images was proposed and implemented. Future work will focus on further refinement and assessment of the technique in the clinical setting. Further research focusing on tumor sites for which there is a clear evidence of benefit from PET/CT-guided RT could possibly help to definitively establish the superiority of tumor delineation based on dynamic PET analysis in comparison to static tumor volume delineation. Evaluation of algorithm performance with images obtained using different tracers, different reconstruction algorithms, and coming from different PET systems could also provide additional insight into the problem. Finally, studies with a larger population of patients featuring patient outcome results and long term follow-up are necessary.

Other applications in clinical oncology such as the assessment of response to treatment are also envisioned. For example, during the course of an effective treatment a rapid decrease of FDG uptake in malignant tissue is expected[27, 108]. This decrease in uptake, which is found to occur prior to any detectable anatomical change with conventional CT images, should allow treatment efficacy to be assessed a short time after treatment commencement. This would provide an opportunity to use the algorithm developed in this work for evaluation of the response of a malignancy to treatment. A clinician monitoring patient progress could potentially make use of tumor volume delineation prior to treatment to determine the therapeutic efficacy. Comparison of tumor volumes delineated on PET scans prior and post treatment may also provide a quantitative measure of tumor response to treatment. The obtained information could prove valuable in

providing an early warning for the need to modify the course of treatment and assure the long term effectiveness of treatment.

Finally, "motion blur" such as would occur for structures affected by respiratory and/or cardiac function and other patient movements may significantly complicate the task of delineation. Patient movement during acquisition could result in discontinuity in the TACs for the affected organ. This issue can be addressed both by PET acquisitions with respiratory gating and post acquisition motion correction. The motion correction of dynamic PET studies can be accomplished through co-registration of all frames against a reference frame or against a time averaged image volume. For example, in a study by *Janssen et al*.[32], frames are co-registered against the last time frame. In this approach, all applied transformations are restricted to rigid body transformations and the normalized cross correlation (or normalized mutual information) serves as cost function. An investigation into the use of alternate motion correction methods is also warranted.

Bibliography

1 R. L. Wahl, Principles and practice of positron emission tomography, edited by Richard L. Wahl, (Lippincott Williams & Wilkins, Philadelphia, 2002), pp. 442.

2 C. B. Caldwell *et al.*, "Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: the impact of 18FDG-hybrid PET fusion," Int.J.Radiat.Oncol.Biol.Phys. 51 (4), 923-931 (2001).

3 R. Bar-Shalom *et al.*, "The additional value of PET/CT over PET in FDG imaging of oesophageal cancer," Eur.J.Nucl.Med.Mol.Imaging 32 (8), 918-924 (2005).

4 J. D. Bradley *et al.*, "Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer," J.Nucl.Med. 45 Suppl 1, 96S-101S (2004).

5 D. Thorwarth and M. Alber, "Implementation of hypoxia imaging into treatment planning and delivery," Radiother.Oncol. 97 (2), 172-175 (2010).

6 D. Thorwarth *et al.*, "Hypoxia dose painting by numbers: a planning study," Int.J.Radiat.Oncol.Biol.Phys. 68 (1), 291-300 (2007).

7 R. M. Pieterman *et al.*, "Preoperative staging of non-small-cell lung cancer with positron-emission tomography," N.Engl.J.Med. 343 (4), 254-261 (2000).

8 H. Zaidi, H. Vees and M. Wissmeyer, "Molecular PET/CT imaging-guided radiation therapy treatment planning," Acad.Radiol. 16 (9), 1108-1133 (2009).

9 T. Pan, L. Chen and C. G. Orton, "Point/counterpoint. PET/CT will become standard practice for radiotherapy simulation and planning," Med.Phys. 35 (9), 3825-3827 (2008).

10 C. Messa *et al.*, "PET/CT and radiotherapy," Q.J.Nucl.Med.Mol.Imag. 50 (1), 4-14 (2006).

11 V. Gregoire *et al.*, "PET-based treatment planning in radiotherapy: a new standard?" J.Nucl.Med. 48 Suppl 1, 68S-77S (2007).

12 L. Xing, J. Siebers and P. Keall, "Computational challenges for image-guided radiation therapy: framework and current research," Semin.Radiat.Oncol. 17 (4), 245-257 (2007).

13 A. C. Paulino, W. L. Thorstad and T. Fox, "Role of fusion in radiotherapy treatment planning," Semin.Nucl.Med. 33 (3), 238-243 (2003).

14 C. Greco *et al.*, "Current status of PET/CT for tumour volume definition in radiotherapy treatment planning for non-small cell lung cancer (NSCLC)," Lung Cancer 57 (2), 125-134 (2007).

15 J. Stroom *et al.*, "Feasibility of pathology-correlated lung imaging for accurate target definition of lung tumors," Int.J.Radiat.Oncol.Biol.Phys. 69 (1), 267-275 (2007).

16 K. Mah *et al.*, "The impact of (18)FDG-PET on target and critical organs in CT-based treatment planning of patients with poorly defined non-small-cell lung carcinoma: a prospective study," Int.J.Radiat.Oncol.Biol.Phys. 52 (2), 339-350 (2002).

17 C. B. Caldwell *et al.*, "Can PET provide the 3D extent of tumor motion for individualized internal target volumes? A phantom study of the limitations of CT and the promise of PET," Int.J.Radiat.Oncol.Biol.Phys. 55 (5), 1381-1393 (2003).

18 C. B. Caldwell *et al.*, "Observer variation in contouring gross tumor volume in patients with poorly defined non-small-cell lung tumors on CT: the impact of 18FDG-hybrid PET fusion," Int.J.Radiat.Oncol.Biol.Phys. 51 (4), 923-931 (2001).

19 J. Bradley *et al.*, "Impact of FDG-PET on radiation therapy volume delineation in non-small-cell lung cancer," Int.J.Radiat.Oncol.Biol.Phys. 59 (1), 78-86 (2004).

20 U. Nestle, S. Kremp and A.-L. Grosu, "Practical integration of [F-18]-FDG-PET and PET-CT in the planning of radiotherapy for non-small cell lung cancer (NSCLC): The technical basis, ICRU-target volumes, problems, perspectives," Radiother.Oncol. 81 (2), 209-225 (2006).

21 R. P. Levy, "PET-CT: evolving role in 3-D radiation therapy," Technol.Cancer.Res.Treat. 5 (2), 101-107 (2006).

22 A. L. Grosu *et al.*, "11C-methionine PET improves the target volume delineation of meningiomas treated with stereotactic fractionated radiotherapy," Int.J.Radiat.Oncol.Biol.Phys. 66 (2), 339-344 (2006).

23 D. E. Heron, S. Beriwal and N. Avril, "FDG-PET and PET/CT in Radiation Therapy Simulation and Management of Patients Who Have Primary and Recurrent Breast Cancer," PET Clinics 1 (1), 39-49 (2006).

24 K. Hamacher, H. H. Coenen and G. Stocklin, "Efficient stereospecific synthesis of no-carrier-added 2-[18F]-fluoro-2-deoxy-D-glucose using aminopolyether supported nucleophilic substitution," J.Nucl.Med. 27 (2), 235-238 (1986).

25 O. Warburg, F. Wind and E. Negelein, "The metabolism of tumors in the body." J.Gen.Physiol. 8 (6), 519-530 (1927).

26 A. Tressaud and G. Haufe, Fluorine and health, edited by A. Tressaud, (Elsevier, 2008), pp. 816.

27 L. A. Drever, "Positron emission tomography target volume delineation", xiv + 134 (2005).

28 J. Skalski, R. L. Wahl and C. R. Meyer, "Comparison of mutual information-based warping accuracy for fusing body CT and PET by 2 methods: CT mapped onto PET emission scan versus CT mapped onto PET transmission scan," J.Nucl.Med. 43 (9), 1184-1187 (2002).

29 P. J. Slomka *et al.*, "Automated 3-dimensional registration of stand-alone (18)F-FDG whole-body PET with CT," J.Nucl.Med. 44 (7), 1156-1167 (2003).

30 W. C. Lavely *et al.*, "Phantom validation of coregistration of PET and CT for image-guided radiotherapy," Med.Phys. 31 (5), 1083-1092 (2004).

31 B. M. Klabbers *et al.*, "Matching PET and CT scans of the head and neck area: development of method and validation," Med.Phys. 29 (10), 2230-2238 (2002).

32 M. H. Janssen *et al.*, "Tumor delineation based on time-activity curve differences assessed with dynamic fluorodeoxyglucose positron emission tomography-computed tomography in rectal cancer patients," Int.J.Radiat.Oncol.Biol.Phys. 73 (2), 456-465 (2009).

33 J. A. Lee, "Segmentation of positron emission tomography images: some recommendations for target delineation in radiation oncology," Radiother.Oncol. 96 (3), 302-307 (2010).

34 A. Teymurazyan *et al.*, "Properties of Noise in Positron Emission Tomography Images Reconstructed with Filtered-Backprojection and Row-Action Maximum Likelihood Algorithm," J.Digital Imaging, 1-12 (2012).

35 A. Teymurazyan *et al.*, "Single Seed Region Growing Algorithm in Dynamic PET Imaging (SSRG/4D-PET) for Tumor Volume Delineation in Radiotherapy Treatment Planning: Theory and Simulation," Nuclear Science, IEEE Transactions on 59 (5), 2020-2032 (2012).

36 S. I. Ziegler, "Positron emission tomography: Principles, technology, and recent developments," Nucl.Phys.A 752, 679C-687C (2005).

37 P. E. Valk, Positron emission tomography, edited by P. E. Valk *et al.* (Springer, London, 2006), pp. 475.

38 G. B. Saha, Basics of PET imaging: physics, chemistry, and regulations, edited by G. B. Saha (Springer, New York, NY, 2005), pp. 206.

39 P. E. Valk, Positron emission tomography: basic science and clinical practice, edited by P. E. Valk *et al* (Springer, London ; New York, 2003), pp. 884.

40 C. S. Levin and E. J. Hoffman, "Calculation of positron range and its effect on the fundamental limit of positron emission tomography system spatial resolution," Phys.Med.Biol. 44 (3), 781-799 (1999).

41 O. Demirkaya, Image Processing with Matlab: Applications in Medicine and Biology, edited by O. Demirkaya *et al.* (Crc Press, Boca Raton, 2008), pp. 441.

42 R. Lecomte, "Novel detector technology for clinical PET," Eur.J.Nucl.Med.Mol.Imaging 36 Suppl 1, S69-85 (2009).

43 S. E .Derenzo, W. W. Moses, "Empirical observation of resolution degradation on positron emission tomographs utilizing block detectors [abstract]," J.Nucl.Med. 34(suppl) (1), 101P (1993).

44 R. Vinke, "Time-of-flight PET with SiPM sensors on monolithic scintillation crystals," KVI - University of Groningen, iii+154 (2011).

45 G. El Fakhri *et al.*, "Quality Assurance for PET and PET/CT Systems," IAEA human health series No. 1, (2009).

46 B. N. Ganguly *et al.*, "Some physical aspects of positron annihilation tomography: A critical review," J.Radioanal.Nucl. 279 (2), 685-698 (2009).

47 C. H. Holdsworth *et al.*, "Investigation of accelerated Monte Carlo techniques for PET simulation and 3D PET scatter correction," IEEE Trans.Nucl.Sci. 48 (1), 74-81 (2001).

48 M. E. Phelps, PET: physics, instrumentation, and scanners, edited by M. E. Phelps (Springer, New York ; Berlin, 2006), pp. 130.

49 P. Razifar, "Novel Approaches for Application of Principal Component Analysis on Dynamic PET Images for Improvement of Image Quality and Clinical Diagnosis," Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology , x + 89 (2005).

50 M. R. Palmer *et al.*, "Transmission Profile Filtering for Positron Emission Tomography," IEEE Trans.Nucl.Sci. 33 (1), 478-481 (1986).

51 P. E. Kinahan and J. G. Rogers, "Analytic 3D Image-Reconstruction using all Detected Events," IEEE Trans.Nucl.Sci. 36 (1), 964-968 (1989).

52 J. A. Thie, "Understanding the standardized uptake value, its methods, and implications for usage," J.Nucl.Med. 45 (9), 1431-1434 (2004).

53 T. Hara *et al.*, "Uptake rates of 18F-fluorodeoxyglucose and 11C-choline in lung cancer and pulmonary tuberculosis: a positron emission tomography study," Chest 124 (3), 893-901 (2003).

54 G. C. Green, "Wavelet-based denoising of cardiac PET data," Carleton University, Ottawa-Carleton Institute for Electrical and Computer Engineering, Department of Systems and Computer Engineering, xiv + 135 (2005).

55 Y. Vardi, L. A. Shepp and L. Kaufman, "A Statistical Model for Positron Emission Tomography," J. Amer. Stat. Assoc. 80 (389), 8-20 (1985).

56 B. M. Tsui *et al.*, "Analysis of recorded image noise in nuclear medicine," Phys.Med.Biol. 26 (5), 883-902 (1981).

57 M. S. Rzeszotarski, "Counting statistics," Radiographics 19 (3), 765-782 (1999).

58 R. W. Rowe and S. Dai, "A pseudo-Poisson noise model for simulation of positron emission tomographic projection data," Med.Phys. 19 (4), 1113-1119 (1992).

59 K. Lange and R. Carson, "EM Reconstruction Algorithms for Emission and Transmission Tomography," J.Comput.Assist.Tomogr. 8 (2), 306-316 (1984).

60 L. A. Shepp and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," Medical Imaging, IEEE Transactions on 1 (2), 113-122 (1982).

61 L. A. Shepp and B. F. Logan, "Fourier Reconstruction of a Head Section," IEEE Trans.Nucl.Sci. NS21 (3), 21-43 (1974).

62 D. J. Kadrmas, "LOR-OSEM: statistical PET reconstruction from raw line-of-response histograms," Phys.Med.Biol. 49 (20), 4731-4744 (2004).

63 J. M. Ollinger and J. A. Fessler, "Positron-emission tomography," Signal Processing Magazine, IEEE 14 (1), 43-55 (1997).

64 P. G. Coxson, R. H. Huesman and L. Borland, "Consequences of using a simplified kinetic model for dynamic PET data," J.Nucl.Med. 38 (4), 660-667 (1997).

65 M. Slifstein, O. R. Mawlawi and M. Laruelle, Chapter 11 (816): Partial volume effect correction: methodological considerations. In: Physiological imaging of the brain with PET, edited by A. Gjedde, S. B. Hansen, G. M. Knudsen and O. B. Paulson, (Academic Press, San Diego, CA, 2000), pp. 413.

66 I. Rodrigues, J. Sanches and J. Bioucas-Dias, "Denoising of Medical Images Corrupted by Poisson Noise," 15th IEEE International Conference on Image Processing 1-5 (ICIP 2008), 1756-1759 (2008).

67 P. Hannequin and J. Mas, "Statistical and heuristic image noise extraction (SHINE): a new method for processing Poisson noise in scintigraphic images," Phys.Med.Biol. 47 (24), 4329-4344 (2002).

68 P. Němeček, "Filtrace šumu ve scintigrafických snímcích metodou založenou na Correspondence Analysis (Noise filtering in scintigraphic images using a method based on Correspondence Analysis.)," Research Reports of CMP, Czech Technical University in Prague, No. 14,, v + 47 (2006).

69 A. Seret, C. Vanhove and M. Defrise, "Resolution improvement and noise reduction in human pinhole SPECT using a multi-ray approach and the SHINE method," Nuklearmedizin 48 (4), 159-165 (2009).

70 T. F. Budinger *et al.*, "Quantitative potentials of dynamic emission computed tomography," J.Nucl.Med. 19 (3), 309-315 (1978).

71 J. Browne and A. B. de Pierro, "A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography," Medical Imaging, IEEE Transactions on 15 (5), 687-699 (1996).

72 M. A. Mandelkern, "Nuclear Techniques for Medical Imaging: Positron Emission Tomography," Annual Review of Nuclear and Particle Science 45, 205-254 (1995).

73 D. W. Wilson and B. M. W. Tsui, "Noise properties of filtered-backprojection and ML-EM reconstructed emission tomographic images," Nuclear Science, IEEE Transactions on 40 (4), 1198-1203 (1993).

74 E. J. Soares, C. L. Byrne and S. J. Glick, "Noise characterization of block-iterative reconstruction algorithms: I. Theory," IEEE Trans.Med.Imaging 19 (4), 261-270 (2000).

75 E. Tanaka and H. Kudo, "Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography," Phys.Med.Biol. 48 (10), 1405-1422 (2003).

76 A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," Journal of the Royal Statistical Society. Series B (Methodological) 39 (1), 1-38 (1977).

77 H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," IEEE Trans.Med.Imaging 13 (4), 601-609 (1994).

78 R. C. Gonzalez and R. E. Woods, Digital image processing, edited by R. C. Gonzalez and R. E. Woods, 3rd ed. (Pearson Prentice Hall, Upper Saddle River, NJ, 2008), pp. 954.

79 NIST/SEMATECH, "e-Handbook of Statistical Methods," (2012) [Online]. Available: http://www.itl.nist.gov/div898/handbook/

80 J. Hilbe, Negative binomial regression, edited by J. Hilbe (Cambridge University Press, Cambridge, 2007), pp. 251.

81 H. H. Barrett, D. W. Wilson and B. M. Tsui, "Noise properties of the EM algorithm: I. Theory," Phys.Med.Biol. 39 (5), 833-846 (1994).

82 F. Pedersen *et al.*, "Principal component analysis of dynamic positron emission tomography images," Eur.J.Nucl.Med. 21 (12), 1285-1292 (1994).

83 J. Portilla *et al.*, "Image denoising using scale mixtures of Gaussians in the wavelet domain," IEEE Trans.Image Process. 12 (11), 1338-1351 (2003).

84 R. C. Gonzalez, R. E. Woods and S. L. Eddins, Digital image processing using MATLAB, edited by R. C. Gonzalez, R. E. Woods and S. L. Eddins (Pearson/Prentice Hall, Upper Saddle River, NJ, 2004), pp. 609.

85 P. Marano, " Application de l'analyse factorielle des correspondances `a la compression des signaux d'images (Application of Factorial Analysis of Correspondances to Compression of Image Signals)," Ann. Télécommunic 27, 163-172 (1972).

86 L. Lebart, Statistique exploratoire multidimensionnelle, edited by L. Lebart, M. Piron and A. Morineau, (Dunod, Paris, 1995), pp. 991.

87 L. Fisher and G. Van Belle, Biostatistics: a methodology for the health sciences, edited by L. Fisher and G. Van Belle (J. Wiley, New York, 1993), pp. 991.

88 M. Malfait and D. Roose, "Wavelet-based image denoising using a Markov random field a priori model," IEEE Trans.Image Process. 6 (4), 549-565 (1997).

89 S. G. Chang, B. Yu and M. Vetterli, Spatially adaptive wavelet thresholding with context modeling for image denoising, edited by S. G. Chang, B. Yu and M. Vetterli (IEEE Computer Soc., Los Alamitos; 10662 Los Vaqueros Circle, P.O. Box 3014, Los Alamitos, CA 90720-1264 USA, 1998), pp. 539.

90 X. Li and M. T. Orchard, Spatially adaptive image denoising under overcomplete expansion, edited by X. Li and M. T. Orchard (IEEE, New York; 345 E 47th St, New York, NY 10017 USA, 2000), pp. 303.

91 A. Pizurica *et al.*, "A joint inter- and intrascale statistical model for Bayesian wavelet based image denoising," IEEE Trans.Image Process. 11 (5), 545-557 (2002).

92 J. L. Starck, D. L. Donoho and E. J. Candes, "Very high quality image restoration by combining wavelets and curvelets," Proc. SPIE 4478, Wavelets: Applications in Signal and Image Processing IX 4478, 9-19 (2001).

93 D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," J.Am.Stat.Assoc. 90 (432), 1200-1224 (1995).

94 H. Zaidi, Quantitative analysis in nuclear medicine imaging, edited by H. Zaidi (Springer Science+Business Media, New York, 2006), pp. 583.

95 R. C. Gonzalez and R. E. Woods, Digital image processing, edited by H. Zaidi 3rd ed. (Pearson Prentice Hall, Upper Saddle River, NJ, 2008), pp. 954.

96 J. Canny, "A Computational Approach to Edge-Detection," IEEE Trans.Pattern Anal.Mach.Intell. 8 (6), 679-698 (1986).

97 N. R. Pal and S. K. Pal, "A Review on Image Segmentation Techniques," Pattern Recognit 26 (9), 1277-1294 (1993).

98 D. E. Heron *et al.*, "Hybrid PET-CT simulation for radiation treatment planning in head-and-neck cancers: a brief technical report," Int.J.Radiat.Oncol.Biol.Phys. 60 (5), 1419-1424 (2004).

99 J. D. Kiffer *et al.*, "The contribution of 18F-fluoro-2-deoxy-glucose positron emission tomographic imaging to radiotherapy planning in lung cancer," Lung Cancer 19 (3), 167-177 (1998).

100 U. Nestle *et al.*, "18F-deoxyglucose positron emission tomography (FDG-PET) for the planning of radiotherapy in lung cancer: high impact in patients with atelectasis," Int.J.Radiat.Oncol.Biol.Phys. 44 (3), 593-597 (1999).

101 U. Nestle *et al.*, "Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-Small cell lung cancer," J.Nucl.Med. 46 (8), 1342-1348 (2005).

102 C. Messa *et al.*, "Feasibility of [18F]FDG-PET and coregistered CT on clinical target volume definition of advanced non-small cell lung cancer," Q.J.Nucl.Med.Mol.Imaging 49 (3), 259-266 (2005).

103 A. C. Riegel *et al.*, "Variability of gross tumor volume delineation in head-and-neck cancer using CT and PET/CT fusion," Int.J.Radiat.Oncol.Biol.Phys. 65 (3), 726-732 (2006).

104 A. van Baardwijk *et al.*, "The current status of FDG-PET in tumour volume definition in radiotherapy treatment planning," Cancer Treat.Rev. 32 (4), 245-260 (2006).

105 J. G. Rajendran *et al.*, "(18)F]FMISO and [(18)F]FDG PET imaging in soft tissue sarcomas: correlation of hypoxia, metabolism and VEGF expression," Eur.J.Nucl.Med.Mol.Imaging 30 (5), 695-704 (2003).

106 D. Thorwarth *et al.*, "A kinetic model for dynamic [18F]-Fmiso PET data to analyse tumour hypoxia," Phys.Med.Biol. 50 (10), 2209-2224 (2005).

107 N. Gupta *et al.*, "Dynamic positron emission tomography with F-18 fluorodeoxyglucose imaging in differentiation of benign from malignant lung/mediastinal lesions," Chest 114 (4), 1105-1111 (1998).

108 P. Rigo *et al.*, "Oncological applications of positron emission tomography with fluorine-18 fluorodeoxyglucose," Eur.J.Nucl.Med. 23 (12), 1641-1674 (1996).

109 A. van Baardwijk *et al.*, "PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumor and involved nodal volumes," Int.J.Radiat.Oncol.Biol.Phys. 68 (3), 771-778 (2007).

110 D. A. Schinagl *et al.*, "Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head and neck cancer," Int.J.Radiat.Oncol.Biol.Phys. 69 (4), 1282-1289 (2007).

111 R. J. Steenbakkers *et al.*, "Observer variation in target volume delineation of lung cancer related to radiation oncologist-computer interaction: a 'Big Brother' evaluation," Radiother.Oncol. 77 (2), 182-190 (2005).

112 R. J. Steenbakkers *et al.*, "Reduction of observer variation using matched CT-PET for lung cancer delineation: a three-dimensional analysis," Int.J.Radiat.Oncol.Biol.Phys. 64 (2), 435-448 (2006).

113 R. Hong *et al.*, "Correlation of PET standard uptake value and CT window-level thresholds for target delineation in CT-based radiation treatment planning," Int.J.Radiat.Oncol.Biol.Phys. 67 (3), 720-726 (2007).

114 A. C. Paulino and P. A. Johnstone, "FDG-PET in radiotherapy treatment planning: Pandora's box?" Int.J.Radiat.Oncol.Biol.Phys. 59 (1), 4-5 (2004).

115 I. F. Ciernik *et al.*, "Radiation treatment planning with an integrated positron emission and computer tomography (PET/CT): a feasibility study," Int.J.Radiat.Oncol.Biol.Phys. 57 (3), 853-863 (2003).

116 E. Brianzoni *et al.*, "Radiotherapy planning: PET/CT scanner performances in the definition of gross tumour volume and clinical target volume," Eur.J.Nucl.Med.Mol.Imaging 32 (12), 1392-1399 (2005).

117 P. Giraud *et al.*, "CT and (18)F-deoxyglucose (FDG) image fusion for optimization of conformal radiotherapy of lung cancers," Int.J.Radiat.Oncol.Biol.Phys. 49 (5), 1249-1257 (2001).

118 T. R. Miller and P. W. Grigsby, "Measurement of tumor volume by PET to evaluate prognosis in patients with advanced cervical cancer treated by radiation therapy," Int.J.Radiat.Oncol.Biol.Phys. 53 (2), 353-359 (2002).

119 A. C. Paulino *et al.*, "Comparison of CT- and FDG-PET-defined gross tumor volume in intensity-modulated radiotherapy for head-and-neck cancer," Int.J.Radiat.Oncol.Biol.Phys. 61 (5), 1385-1392 (2005).

120 J. F. Daisne *et al.*, "Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios: influence of reconstruction algorithms," Radiother.Oncol. 69 (3), 247-250 (2003).

121 D. Hellwig *et al.*, "18F-FDG PET for mediastinal staging of lung cancer: which SUV threshold makes sense?" J.Nucl.Med. 48 (11), 1761-1766 (2007).

122 U. Nestle *et al.*, "Target volume definition for 18F-FDG PET-positive lymph nodes in radiotherapy of patients with non-small cell lung cancer," Eur.J.Nucl.Med.Mol.Imaging 34 (4), 453-462 (2007).

123 N. Otsu, "Threshold Selection Method from Gray-Level Histograms," IEEE Trans.Syst.Man Cybern. 9 (1), 62-66 (1979).

124 S. S. Reddi, S. F. Rudin and H. R. Keshavan, "An Optimal Multiple Threshold Scheme for Image Segmentation," IEEE Trans.Syst.Man Cybern. 14 (4), 661-665 (1984).

125 L. Drever *et al.*, "Iterative threshold segmentation for PET target volume delineation," Med.Phys. 34 (4), 1253-1265 (2007).

126 L. M. Le Cam and J. Neyman, "Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, held at the Statistical Laboratory, University of California, June 21-July 18, 1965 and December 27, 1965-January 7, 1966," Biology and problems of health, Vol. IV, pp. 934 (1967).

127 G. B. Coleman and H. C. Andrews, "Image Segmentation by Clustering," Proc IEEE 67 (5), 773-785 (1979).

128 J. C. Bezdek, Pattern recognition with fuzzy objective function algorithms, edited by J. C. Bezdek (Plenum Press, New York, 1981), pp. 256.

129 A. E. Boudraa *et al.*, "Delineation and quantitation of brain lesions by fuzzy clustering in positron emission tomography," Comput.Med.Imaging Graph. 20 (1), 31-41 (1996).

130 A. E. Boudraa *et al.*, "Left-Ventricle Automated Detection Method in Gated Isotopic Ventriculography using Fuzzy Clustering," IEEE Trans.Med.Imaging 12 (3), 451-465 (1993).

131 A. E. Boudraa *et al.*, "Automated detection of the left ventricular region in gated nuclear cardiac imaging," IEEE Trans.Biomed.Eng. 43 (4), 430-437 (1996).

132 H. Zaidi *et al.*, "Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging," Phys.Med.Biol. 47 (7), 1143-1160 (2002).

133 K. J. Van Laere *et al.*, "Analysis of clinical brain SPECT data based on anatomic standardization and reference to normal data: An ROC-based comparison of visual, semiquantitative, and voxel-based methods," J.Nucl.Med. 43 (4), 458-469 (2002).

134 P. J. Slomka *et al.*, "Three-dimensional demarcation of perfusion zones corresponding to specific coronary arteries: application for automated interpretation of myocardial SPECT," J.Nucl.Med. 36 (11), 2120-2126 (1995).

135 S. Z. Li, Markov Random Field Modeling in Image Analysis, edited by Stan Z. Li, (Springer-Verlag, London, 2009), pp. 357.

136 M. Skurichina and RPW Duin, "Bagging for linear classifiers," Pattern Recognit 31 (7), 909-930 (1998).

137 S. Raudys and RPW Duin, "Expected classification error of the Fisher linear classifier with pseudo-inverse covariance matrix," Pattern Recog.Lett. 19 (5-6), 385-392 (1998).

138 Y. Hamamoto, Y. Fujimoto and S. Tomita, "On the estimation of a covariance matrix in designing Parzen classifiers," Pattern Recognit 29 (10), 1751-1759 (1996).

139 V. N. Vapnik, The nature of statistical learning theory, edited by V. N. Vapnik (Springer, New York, 1995), pp. 188.

140 L. A. Drever *et al.*, "Comparison of three image segmentation techniques for target volume delineation in positron emission tomography," J.Appl.Clin.Med.Phys. 8 (2), 93-109 (2007).

141 M. Kass, A. Witkin and D. Terzopoulos, "Snakes - Active Contour Models," Int.J.Comput.Vis. 1 (4), 321-331 (1987).

142 H. Guo *et al.*, "Clustering huge data sets for parametric PET imaging," BioSystems 71 (1-2), 81-92 (2003).

143 K. P. Wong *et al.*, "Segmentation of dynamic PET images using cluster analysis," IEEE Trans.Nucl.Sci. 49 (1), 200-207 (2002).

144 P. D. Acton *et al.*, "Automatic segmentation of dynamic neuroreceptor single-photon emission tomography images using fuzzy clustering," Eur.J.Nucl.Med. 26 (6), 581-590 (1999).

145 Y. Kimura, M. Senda and N. M. Alpert, "Fast formation of statistically reliable FDG parametric images based on clustering and principal components," Phys.Med.Biol. 47 (3), 455-468 (2002).

146 J. G. Brankov *et al.*, "Segmentation of dynamic PET or fMRI images based on a similarity metric," IEEE Trans.Nucl.Sci. 50 (5), 1410-1414 (2003).

147 M. E. Kamasak and B. Bayraktar, "Clustering dynamic PET images on the projection domain," IEEE Trans.Nucl.Sci. 54 (3), 496-503 (2007).

148 J. Kim *et al.*, "Segmentation of VOI from multidimensional dynamic PET images by integrating spatial and temporal features," IEEE Trans.Inf.Technol.Biomed. 10 (4), 637-646 (2006).

149 X. Li, D. Feng and K. Chen, "Optimal image sampling schedule: a new effective way to reduce dynamic image storage space and functional image processing time," IEEE Trans.Med.Imaging 15 (5), 710-719 (1996).

150 A. M. Spence *et al.*, "18F-FDG PET of gliomas at delayed intervals: improved distinction between tumor and normal gray matter," J.Nucl.Med. 45 (10), 1653-1659 (2004).

151 C. J. Kelly and M. Brady, "A model to simulate tumour oxygenation and dynamic [18F]-Fmiso PET data," Phys.Med.Biol. 51 (22), 5859-5873 (2006).

152 D. Thorwarth *et al.*, "Kinetic analysis of dynamic 18F-fluoromisonidazole PET correlates with radiation treatment outcome in head-and-neck cancer," BMC Cancer 5, 152 (2005).

153 Y. Kimura *et al.*, "MAP-based kinetic analysis for voxel-by-voxel compartment model estimation: detailed imaging of the cerebral glucose metabolism using FDG," Neuroimage 29 (4), 1203-1211 (2006).

154 L. G. Strauss *et al.*, "18F-FDG kinetics and gene expression in giant cell tumors," J.Nucl.Med. 45 (9), 1528-1535 (2004).

155 E. U. Nitzsche *et al.*, "Non-invasive differentiation of pancreatic lesions: is analysis of FDG kinetics superior to semiquantitative uptake value analysis?" Eur.J.Nucl.Med.Mol.Imaging 29 (2), 237-242 (2002).

156 L. G. Strauss *et al.*, "Assessment of quantitative FDG PET data in primary colorectal tumours: which parameters are important with respect to tumour detection?" Eur.J.Nucl.Med.Mol.Imaging 34 (6), 868-877 (2007).

157 A. Dimitrakopoulou-Strauss, L. G. Strauss and J. Rudi, "PET-FDG as predictor of therapy response in patients with colorectal carcinoma," Q.J.Nucl.Med. 47 (1), 8-13 (2003).

158 K. Levenberg, "A Method for the Solution of Certain Problems in Least-Squares," Q Appl Math 2 (2), 431-441 (1944).

159 D. W. Marquardt, "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," Journal of the Society for Industrial and Applied Mathematics 11 (2), 431-441 (1963).

160 G. A. Watson, Numerical analysis : proceedings of the biennial conference : held at Dundee, June 28-July 1, 1977, edited by G. A. Watson (Springer-Verlag, Berlin ; New York, 1978), pp. 199.

161 T. W. Ridler and S. Calvard, "Picture Thresholding using an Iterative Selection Method," IEEE Trans.Syst.Man Cybern. 8 (8), 630-632 (1978).

162 L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," IEEE Trans.Image Process. 2 (2), 176-201 (1993).

163 T. Torizuka *et al.*, "Short dynamic FDG-PET imaging protocol for patients with lung cancer," Eur.J.Nucl.Med. 27 (10), 1538-1542 (2000).

164 M. Bentourkia and H. Zaidi, "Tracer Kinetic Modeling in PET," PET Clin 2 (2), 267-277 (2007).

165 R. Jia *et al.*, "Optimal dual time point for FDG-PET in the differentiation of benign from malignant lung lesions: a simulation study," Conf.Proc.IEEE Eng.Med.Biol.Soc. 2007, 4169-4172 (2007).

166 O. Selberg *et al.*, "Use of positron emission tomography for the assessment of skeletal muscle glucose metabolism," Nutrition 18 (4), 323-328 (2002).

167 H. Qiao *et al.*, "Method for tumor recognition with short dynamic PET images: Theory and simulation study," Prog.Nat.Sci. 19 (1), 73-77 (2009).

168 D. Feng *et al.*, "A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: theory and simulation study," IEEE Trans.Inf.Technol.Biomed. 1 (4), 243-254 (1997).

169 D. Feng, X. Li and S. C. Huang, "A new double modeling approach for dynamic cardiac PET studies using noise and spillover contaminated LV measurements," IEEE Trans.Biomed.Eng. 43 (3), 319-327 (1996).

170 D. Feng, S. C. Huang and X. Wang, "Models for computer simulation studies of input functions for tracer kinetic modeling with positron emission tomography," Int.J.Biomed.Comput. 32 (2), 95-110 (1993).

171 E. Høgh-Rasmussen, "BBTools - a Matlab Toolbox for Black-Box Computations," Neurobiology Research Unit, Copenhagen University Hospital, 2005 [Online]. Available: http://nru.dk/software/bbtools/

172 H. M. Wu *et al.*, "A modeling method to improve quantitation of fluorodeoxyglucose uptake in heterogeneous tumor tissue," J.Nucl.Med. 36 (2), 297-306 (1995).

173 M. E. Phelps *et al.*, "Tomographic measurement of local cerebral glucose metabolic rate in humans with (F-18)2-fluoro-2-deoxy-D-glucose: validation of method," Ann.Neurol. 6 (5), 371-388 (1979).

174 X. Geets *et al.*, "A gradient-based method for segmenting FDG-PET images: methodology and validation," Eur.J.Nucl.Med.Mol.Imaging 34 (9), 1427-1438 (2007).

175 R. M. Pieterman *et al.*, "Comparison of (11)C-choline and (18)F-FDG PET in primary diagnosis and staging of patients with thoracic cancer," J.Nucl.Med. 43 (2), 167-172 (2002).