Wavelet-based Denoising

of

Cardiac PET Data

by

Geoffrey C. Green

A thesis submitted to

The Faculty of Graduate Studies and Research

In partial fulfillment of the requirements for the degree of

Master of Applied Science

in Electrical Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario, Canada

January 2005

The undersigned hereby recommend to

the Faculty of Graduate Studies and Research

acceptance of the thesis

Wavelet-based Denoising

of

Cardiac PET Data

Submitted by

Geoffrey C. Green,

in partial fulfillment of the requirements

for the degree of Master of Applied Science

_____

Dr. Aysegul Cuhadar, Thesis Supervisor

_____

Dr. Robert deKemp, Thesis Supervisor

_____

Chair, Department of Systems and Computer Engineering

Carleton University

January 2005

To my late father, who taught me the importance of education.

To my mother, whose strength and positive attitude toward life inspires me.

And to Katherine and Timothy, my new family.

Acknowledgments

I would like to thank my thesis supervisor, Dr. Aysegul Cuhadar, for giving me the chance to explore a research career in biomedical engineering. Her continual support, technical insight, and dedication to this research has made this opportunity most enjoyable.

I would also like to thank Dr. Rob deKemp at the Ottawa Heart Institute for being so accessible and helpful, for the guidance he provided during this research project, and for taking the time to explain the practical and clinical aspects of PET imaging.

Abstract

This thesis focuses on denoising of positron emission tomography (PET) data. Cardiac PET scans generated using a rubidium-82 radiotracer are a convenient, non–invasive method of diagnosing heart disease, but suffer from a high degree of noise. Denoising methods based on the wavelet transform are capable of outperforming existing clinical methods due to their ability to better preserve detail while simultaneously suppressing noise at multiple scales. We investigate the applicability of recently developed wavelet denoising methods to cardiac PET data. A comprehensive set of experiments is performed, in which combinations of these techniques are applied to the different decomposition levels of wavelet coefficients. By doing so, we determine the relevant importance of each (and the domain in which it is applied) to the overall quality of the denoised result. With this information, we propose PET denoising protocols that substantially improve image quality (for static studies) and lead to better measures of myocardial perfusion (for dynamic studies).

Contents

List of Tables

List of Figures

List of Acronyms

^{82}Rb	Rubidium-82
AWGN	Additive white Gaussian noise
CAD	Coronary artery disease
CSR	Cross scale regularization
CT	Computed tomography
CWT	Continuous wavelet transform
DDWT	Discrete dyadic wavelet transform
DSP	Digital signal processing
DWT	Discrete wavelet transform
ECG	Electrocardiogram
FBP	Filtered backprojection
FWHM	Full width at half maximum
FWT	Fast wavelet transform
GLPF	Gaussian low pass filter
IDDWT	Inverse discrete dyadic wavelet transform

IDWT	Inverse discrete wavelet transform
iid	Independent and identically distributed
LOR	Line of response
LV	Left ventricle
MLEM	Maximum likelihood expectation maximization
MMSE	Minimum mean squared error
MRA	Multiresolution analysis
MRI	Magnetic resonance imaging
MSE	Mean squared error
OHI	Ottawa Heart Institute
OSEM	Order subset expectation maximization
PDF	Probability density function
PET	Positron emission tomography
PSNR	Peak signal to noise ratio
Rb	Rubidium
ROI	Region of interest
SA	Short axis, spatially adaptive
TAC	Time activity curve
TBV	Total blood volume
TI	Translation invariant

Chapter 1

Introduction

1.1 Motivation

Positron emission tomography (PET) is a nuclear medicine imaging modality that is used to observe and measure physiological processes as they occur in the human body (*in vivo*). This is accomplished by injecting a positron emitting radiotracer into a patient, which is selectively taken up in the organ(s) of interest. As the tracer decays, the resulting radioactivity is measured by the detectors of a PET scanner. These data are then used to reconstruct either static or time-varying images that attempt to measure, as closely as possible, the tracer distribution present in the body. The resulting set of images, and the quantitative information that can be extracted from them, provides a physician with a representation of underlying pathology. This is impossible with other modalities that provide anatomical information only, for instance computed tomography (CT) and conventional magnetic resonance imaging (MRI). PET is employed in a diverse variety of clinical applications, among these the diagnosis of psychiatric disorders (such as Alzheimer's disease) and the measurement of the extent of cancerous tissue [1].

In this thesis, we are interested in an application currently being used at the Na-

tional Cardiac PET Center, University of Ottawa Heart Institute (OHI) - the measurement of blood flow to heart muscle tissue (termed *myocardial perfusion*). Cardiac PET scans provide useful measures of the location and extent of damaged heart tissue, which can be caused by coronary artery disease (CAD). This information is used by a cardiologist to prescribe an appropriate course of treatment for patients with CAD. For example, in order to decide whether or not bypass surgery is potentially helpful, the viability of the myocardium must be established. If the tissue is damaged beyond repair (a diagnosis possible with a PET study), there is no point in pursuing an expensive surgical procedure[2].
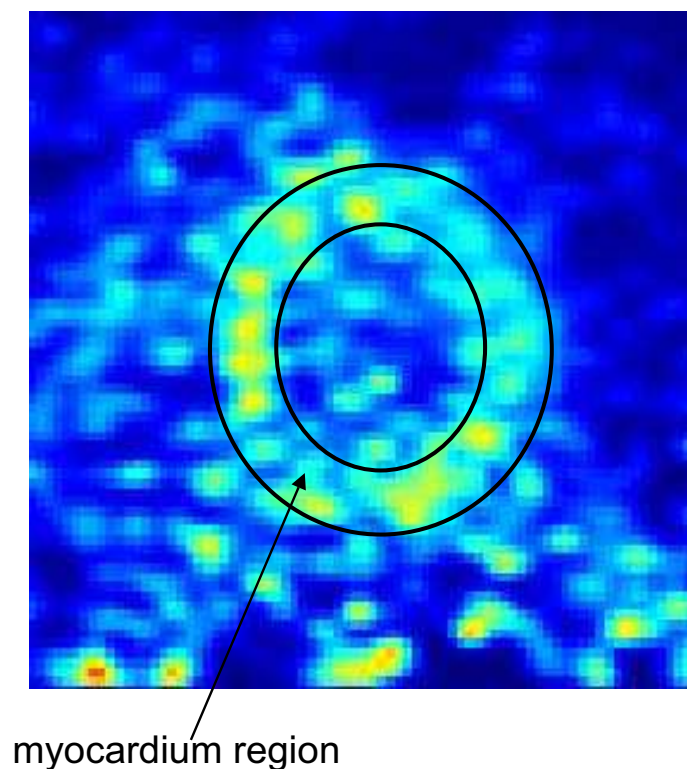
Figure 1.1: A enlarged view of a short-axis oriented image of the heart muscle generated with a ^{82}Rb PET scan. There is a large amount of noise obscuring the true "ring-shaped" myocardium region (outlined with black lines).

Rubidium-82 (^{82}Rb) is a suitable radiotracer choice for perfusion studies. It is generator produced and thus does not require an on–site cyclotron for its production.

^{82}Rb is analogous to potassium in its affinity for muscle tissue and it does not affect normal metabolism, so it is well incorporated into the cardiac uptake process. Its short physical half-life (76s) allows repeated studies within a short time period, allowing flexibility when designing clinical studies. Unfortunately, ^{82}Rb PET scans produce data highly contaminated by noise, as is evident in Figure 1.1. As a result, dynamic curves of radiotracer activity that are derived from these data are error-prone and the resulting perfusion estimates are usually biased[3, 4].

It is desirable to perform image processing techniques in order to suppress the noise in these images. Such noise-reduced images would provide more accurate indications of cardiac function and lessen the chance of misdiagnoses and inappropriate treatment due to poor image quality. Conventional approaches to noise reduction in cardiac PET (which are still used in practice) include spatial convolution in the image domain with a Gaussian kernel [5]. This is essentially a low pass filtering operation that, though satisfactory in certain cases, tends to oversmooth small-scale image features for a given desired resolution. This is especially detrimental when precise quantitative information is extracted.

1.2   Thesis Objectives

The goal of this thesis is to develop noise reduction (or *denoising*) methods that improve the quality of cardiac ^{82}Rb PET scans, and illustrate their effectiveness and robustness when used to measure myocardial perfusion.

The denoising methods we propose are based on a wavelet representation. Noise reduction using the wavelet transform has the potential to outperform existing methods for medical images, as it exploits the wavelet's natural ability to separate signal from noise at *multiple image scales.* Cardiac PET images are characterized by slowly changing uniform uptake regions (*e.g.* in the healthy myocardium) that might be "in-

terrupted" by smaller scale discontinuities (*e.g.* due to a cardiac defect). A wavelet decomposition of a signal is based on a collection of compactly supported oscillatory basis functions that are related to each other through scalings and translations. The inclusion of localized fine–scale functions in the basis allows one to better reconstruct diagnostically significant details at multiple scales. Compared to conventional denoising techniques, this provides an inherent advantage when dealing with non-stationary signals (like cardiac PET images) and can thereby extend the current state of the art in cardiac PET denoising.

It is anticipated that applying wavelet-based denoising methods to cardiac PET data will lead to higher quality images and more accurate measures of myocardial perfusion. Evaluating the success of noise reduction in a clinical setting is not trivial, given the lack of *a priori* knowledge of the "true" radiotracer distribution hidden in the body. In this thesis, the denoising performance is assessed in the following ways:

1. the use of phantom data modeling the anatomical regions under study,

2. the use of performance parameters extracted from the tracer kinetic model, and

3. a subjective visual assessment by qualified PET personnel.

1.3   Thesis Contributions

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

- two state of the art methods in wavelet denoising have been adopted for multi-dimensional ^{82}Rb cardiac PET scans – *spatially adaptive thresholding* and *cross scale regularization*. We investigate the relative effect that these methods have on the denoised result when applied in combination across multiple scales. In addition, we perform denoising in multiple domains, and evaluate the extent to which noise reduction in each domain is important.

- a set of novel denoising protocols are proposed that increase the quality of static PET scans and lead to better measures of myocardial perfusion. These are a hybrid of the two aforementioned methods, performing cross scale regularization at the finest scales, in combination with spatially adaptive thresholding.

- a comparison between denoising PET images with the clinical protocol and the proposed methods is performed.

- a graphical user interface is developed in Matlab to facilitate the evaluation of denoising results, and to provide a framework for further study of wavelet-based denoising on cardiac PET data.

1.4 Thesis Outline

In Chapter 2, we present background information on PET. The basic physical process of positron emission and photon generation is introduced, followed by a description of the PET scanner's detection mechanism. We discuss image reconstruction, which converts the raw measurements in the projection domain to a time series of three-dimensional images. Since our focus in this work is on noise reduction, we conclude the chapter with a section on the main sources of noise in PET.

Chapter 3 describes the PET application of interest in this thesis. We discuss basic cardiac anatomy, cardiac function and the relevant spatial and spatiotemporal representations of the heart. This is followed by a description of both static and dynamic PET studies, concluding with a brief overview of the physiological model used with ^{82}Rb to extract flow measures.

In Chapter 4, we switch our attention to the field of image denoising. We first present classical noise reduction techniques and discuss their limitations. We then motivate how these shortcomings can be overcome with a wavelet-based representation. A review of the wavelet denoising literature is presented, including the methods

most promising for cardiac PET data. The chapter is concluded by a discussion of the discrete dyadic wavelet transform (DDWT) - both the motivation for using it and the necessary mathematical background.

Chapter 5 presents the application of wavelet-based denoising methods to cardiac PET data. After describing the input data and examining its relevant noise properties, we provide details on an efficient filter bank implementation of the DDWT. Following this, we give a full description of the denoising protocols and establish the various figures of merit that are used to measure denoising performance. The results are presented, after which we summarize the conclusions that we arrive at as a result of this study.

Lastly, Chapter 6 summarizes the thesis and discusses future work directions that would further the value of our research.

Chapter 2

Background — PET

In this chapter, we present a brief overview of the PET image formation process. The process of positron emission is introduced in Section 2.1, followed by a description of the PET scanner and sinogram arrays that detect and record these emissions in Sections 2.2 and 2.3 respectively. Section 2.4 describes the most clinically popular methods of reconstructing images from these projections – filtered backprojection (FBP) and ordered subset expectation-maximization (OSEM). We conclude this chapter with a discussion of the major noise sources in PET and how this noise degrades the images in Section 2.5.

2.1 Physics of Positron Emission

For the reasons outlined in the previous chapter, ^{82}Rb is a suitable radiotracer for cardiac PET studies. It starts to appear in the myocardium within minutes after it has been administered. Being an unstable isotope, the nucleus decays according to the following process:

$$^{82}\text{Rb} \longrightarrow ^{82}\text{Kr} + \beta^+ + \nu \tag{2.1}$$

This results in the appearance of a fast-moving positron (β^+, an anti-matter electron) in the myocardium, as well as a neutrino ν. After traveling a short distance (a random quantity called the *positron range*, approximately 3 mm for ^{82}Rb in heart tissue), the positron combines with a neighbouring electron. The collision between the positron and electron (β^-), antiparticles of one another, causes an *annihilation*, which results in the conversion of their mass to energy:

$$\beta^+ + \beta^- \longrightarrow 2\gamma \tag{2.2}$$

What remains is a pair of gamma-ray photons (γ) (each of energy 511 keV) that travel away from the annihilation site at almost exactly 180 degrees apart. It is these photons that escape the body and can be subsequently detected by the PET scanner [1]. The positron emission and annihilation process is shown in Figure 2.1.
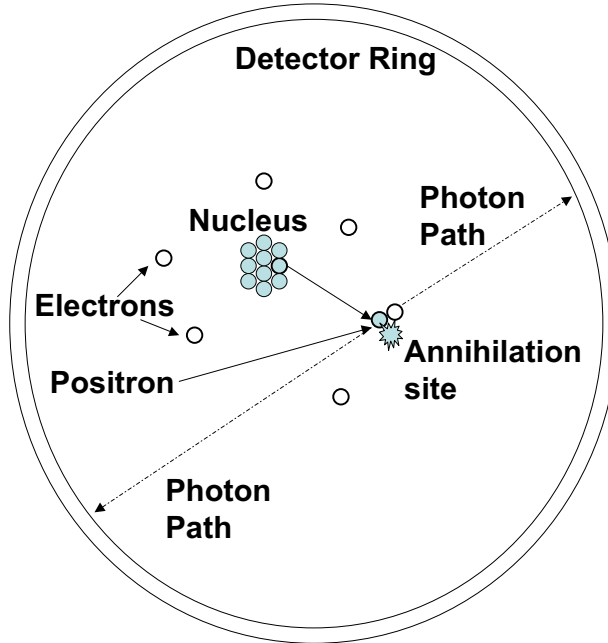
Figure 2.1: Positron emission from ^{82}Rb nucleus followed by positron / electron annihilation event generating gamma-ray photons, after positron travels a distance equal to the positron range.

2.2   Description of PET Scanner and Counts

The ECAT ART PET scanner (CTI Systems, Knoxville, TN) [6] is used at OHI and is shown in Figure 2.2, along with the co-ordinate system axes used in subsequent discussion. The scanner consists of a gantry on which a patient, who has had an injected radiotracer dose, lies prone during the duration of the PET scan. This gantry can slide along the z direction, so that the appropriate part of the body can be positioned in the detector aperture, which consists of 24 concentric rings of detector arrays surrounding the patient. Each ring (in the $x - y$ plane) consists of a series of small individual detectors oriented around the circumference of the scanner aperture. When an annihilation occurs, the resulting photons leave the patient's body in opposite directions. Under certain conditions (see below), they can be used to locate a positron emission location - these events are termed *true counts*. The line connecting the two detectors that counted the photons is called the *line of response* (LOR). To distinguish these from other types of counts, the arrival times of the photons at each detector are recorded and compared to see if they lie within the same time window[1]. If so, such photons are said to be *in coincidence* and are assumed to originate from the same annihilation [5].

True counts are recorded when the photons do not get scattered by tissue encountered on its exit path from the body. When scattering occurs, one or both photons are "knocked off" their original (almost) straight line trajectory. They can still arrive at different detectors within the coincidence time window, but the detector has no way of knowing that one or both of the photons have arrived along a scattered path. In this case, the straight line between detectors will not correspond to the line on which the actual annihilation occurred, so the resulting count (called a *scattered count*) is mispositioned. Another source of error occurs with *random counts*, whereby photons

---

[1]This window duration is typically on the order of several nanoseconds (10^{-9} s), since the photons travel at the speed of light, 3×10^8 m/s, and the distance traveled is typically $10 - 20$ cm.
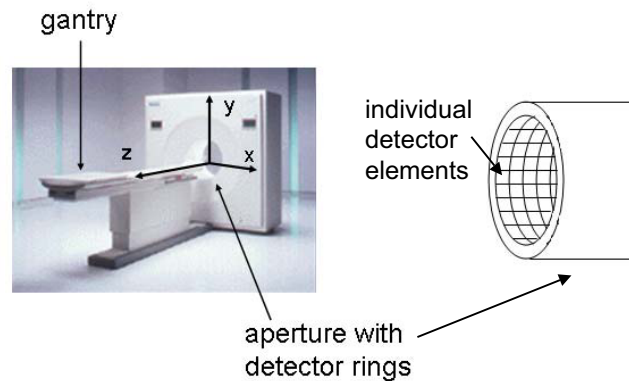
Figure 2.2: ECAT ART PET scanner, with relevant components iden-
tified. The coordinate system axes used in this thesis is
included.

not due to the same annihilation fall within the coincidence timing window. This also
results in an incorrectly computed LOR. As expected, the effect of random counts is
more pronounced during scans of very high activity [1]. Figure 2.3 illustrates these
three types of counts.

It is also important to note that, due to the finite field of view of the PET scanner,
not all positron annihilations will produce two photons that hit the detectors. In this
category, we have:

- those events in which neither of the annihilation photons get detected, and

- those events in which one photon gets detected, but the other does not. In this
  case, the coincidence timing window will expire and the detected photon will
  be discarded.

In addition, it is possible that a photon arrives at a detector, but before the coinci-
dence window expires, yet another photon arrives at the same detector. In this case,
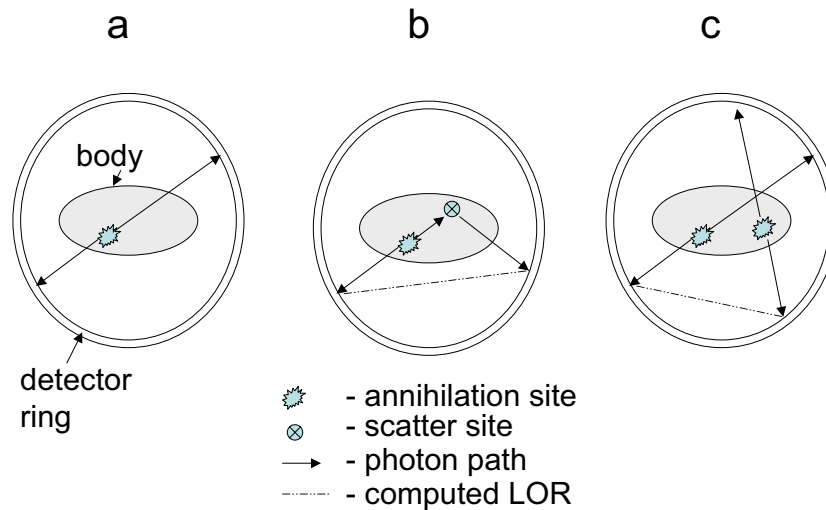
Figure 2.3: Illustration of (a) a true count, (b) a scattered count, and
(c) a random count. In each case, the photons are detected
within the same coincidence timing window.

since the detector can only identify and maintain coincidence timing for one photon
at a time, both photons are discarded. This time interval in which the detector is
unavailable is called *dead time* and is a function of count rate and detector size. High
dead times result in lower counts, and hence a lower signal to noise ratio.

2.3   PET Measurements and Sinograms

The raw data for a PET scan consist of radioactive counts described above, col-
lected over a suitably long time interval. As the scan proceeds and annihilations
occur, each count is stored in an array in memory. The image has not yet been
reconstructed at this stage, so this storage array does not correspond to the 2D or
3D image representation of radiotracer activity. Instead, a count is specified by its
LOR only - the actual location (*i.e.* x, y, z coordinates) of the annihilation along this

LOR is unknown[2]. Because there are thousands of individual detectors, the number of combinations of possible pairs of detectors (and therefore the number of LORs) is in the millions.
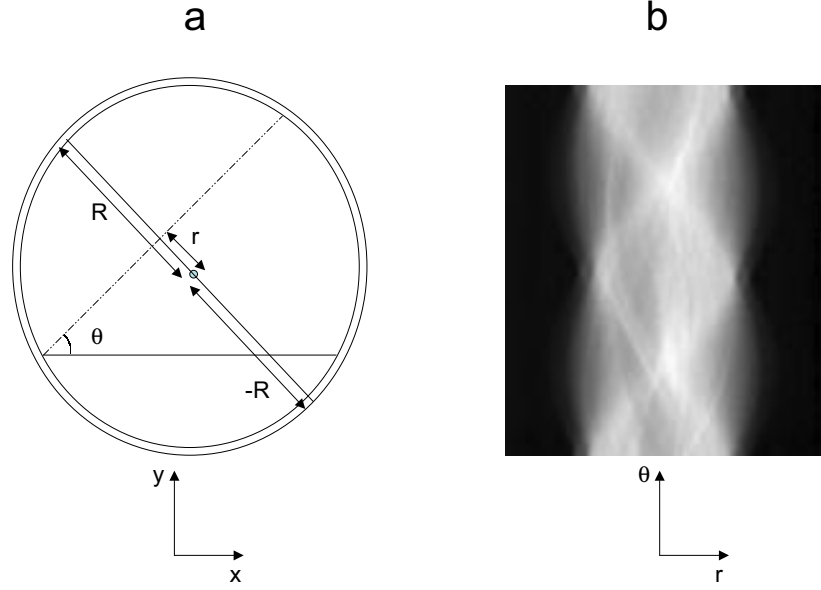
Figure 2.4: (a) Line of Response in a PET scanner aperture, (b) Sinogram array

As seen in Figure 2.4 for a single ring of detectors (*i.e.* fixed z value), a LOR is specified by two parameters - its distance from the center of the aperture (r) and the angle (θ) that the normal to the LOR makes with the $+x$ axis. The memory locations that holds the raw data are indexed by these parameters in a two-dimensional array called a *sinogram*[3]. Each point (r, θ) in the sinogram specifies a LOR and the integer value stored at that location is the number of counts that has occurred along that LOR during the scan time. The data stored in a horizontal row of the sinogram (*i.e.*

---

[2]Calculating the annihilation location based on the difference between photon arrival times at the detectors forms the basis of another type of emission tomography, called time-of-flight (TOF) PET.

[3]The term "sinogram" is used because a) the counts resulting from a point source in the scanner's field of view will trace a **sinusoidal** pattern in the 2D array, and b) the array is accumulating counts into bins, similar to a **histogram**.

fixed θ, varying r) correspond to the counts that occur along all parallel LORs in the aperture at that θ value - these counts make up a *projection*.

Finally, we note that the OHI PET scanner used in this study operates in 3D mode. There are no extended septa separating the rings of detectors, but the count data are rebinned into a set of 2D sinograms, one for each slice (z value). Image reconstruction (explained in the next section) is performed in two dimensions on a *per-slice* basis, based on the 2D sinogram data at a single z value. In order to ensure practical scan times, three-dimensional image reconstruction is not used typically[1].

2.4   Image Reconstruction from Projections

The preceding section described how the raw PET data (radioactive counts) are stored. The next step is to use these *projection domain* data to generate images that show the spatial distribution of ^{82}Rb in the heart muscle as accurately as possible. The problem of image reconstruction from projections was pioneered by J. Radon in 1917 and has undergone successive refinements since then, in such diverse areas of application as radio astronomy, electron microscopy, and the field at hand, medical imaging[7].

As mentioned, the photon activity that occurs along a LOR is captured by the pair of detectors at the endpoints of that line. The counts recorded by the detectors correspond to an integration of the activity along the corresponding LOR over time. This process results in a loss of spatial information - namely, the location of individual annihilation sites along that LOR.

The configuration in two dimensions is shown in Figure 2.5. We denote the actual radiotracer distribution being imaged in the heart region as $f(x, y)$, f being a positive, integer valued function measured in units of Bq/cc.[4] For now, we assume that x and
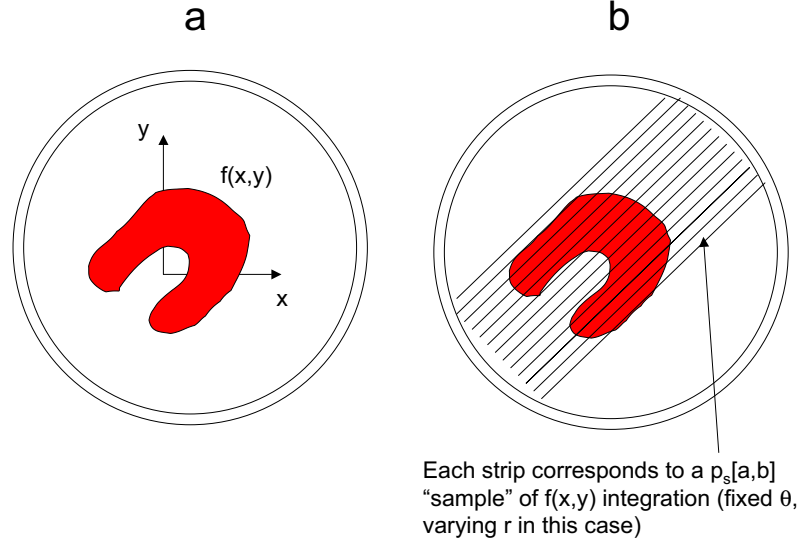
Figure 2.5: (a) Two-dimensional radiotracer distribution $f(x,y)$ in the myocardium, as positioned within the detector ring of a PET scanner, (b) "thin-strip" approximation to line integral.

y are continuous spatial variables.

The process of projecting the radiotracer concentration $f(x,y)$ onto the detector rings is described mathematically with the continuous version of the Radon transform operator $\mathcal{R}_c$ that produces the projection data $p(\theta, r)$:

$$p(\theta, r) = \mathcal{R}_c \{f(x,y)\}$$
$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\delta(r - x\cos\theta - y\sin\theta)dxdy \tag{2.3}$$

where r is a continuous radial variable defined on $[-R, R]$, θ is a continuous angle variable defined on $[0, \pi]$, and $\delta(r)$ corresponds to the continuous-time Dirac delta function.

To recover the image from its (continuous) projection representation, one uses the

---

[4] The unit Bq (becquerel) corresponds to one nuclear decay per second, and cc is a cubic centimeter

continuous version of the inverse Radon transform operator, $\mathcal{R}_c^{-1}$:

$$
\begin{aligned}
f(x,y) &= \mathcal{R}_c^{-1}\left\{p(\theta,r)\right\} \\
&= \frac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^\infty P_\theta(\omega) e^{j\omega(x\cos\theta + y\sin\theta)} R(\omega)\, d\omega\, d\theta
\end{aligned}
\tag{2.4}
$$

where $P_\theta(\omega)$ is the 1D continuous time Fourier transform (in the radial variable) of $p(\theta, r)$, and $R(\omega) = |\omega|$ is the ramp filter, specified in the frequency domain. This relation is based on the following theorem, illustrated in Figure 2.6.

Fourier Slice Theorem. *The Fourier Slice Theorem states that the Fourier transform of a projection of a two-dimensional function $f(x,y)$, seen from an angle θ, equals the slice of the 2D Fourier Transform of $f(x,y)$, $\mathcal{F}\{f(x,y)\} = F(\omega_x, \omega_y)$, under that same angle θ [7].*
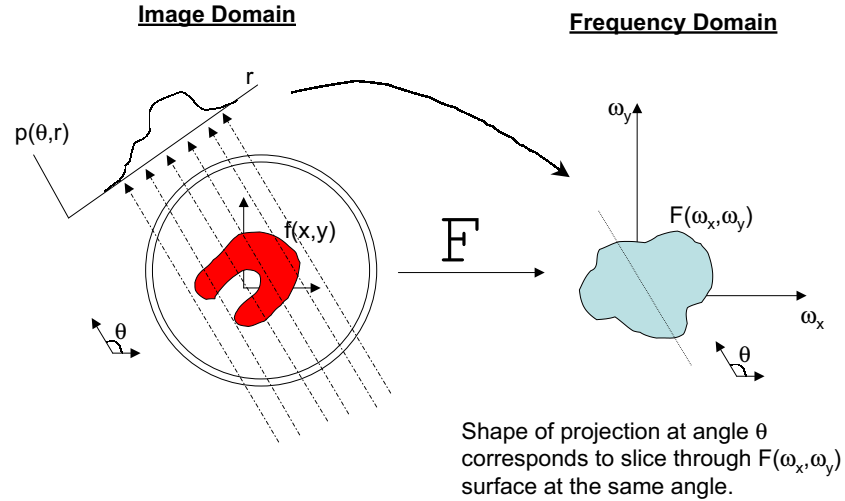
Figure 2.6: Illustration of the Fourier Slice Theorem

We emphasize that the ideal $\mathcal{R}_c^{-1}$ operator in equation 2.4 will only give the "true" tracer distribution $f(x,y)$ when the input projection data are *noise free*. In reality, the

noise processes in PET scanners contaminate the measured projection data, adding an undesired noise contribution in the observed image. For reasons previously outlined, a noise-free image is neither physically possible, nor a reasonable assumption.

In practice, a PET scanner is a digital device, so the continuously defined quantities in equations 2.3 and 2.4 pose a practical problem for implementation. A discrete version of the reconstructed image, which we call f_s, is obtained from the continuous projection data with a sampled version of the inverse Radon transform operator, $\mathcal{R}_s^{-1}$:

$$
\begin{aligned}
f_s[m,n] &= \mathcal{R}_s^{-1}\{p(\theta, r)\} \\
&= \frac{1}{4\pi^2} \int_0^\pi \int_{-\infty}^\infty P_\theta(\omega) e^{j\omega(m\cos\theta + n\sin\theta)} R(\omega) d\omega d\theta
\end{aligned}
\tag{2.5}
$$

where $m \in [1, \ldots, M]$ is the sample number in the x direction, and $n \in [1, \ldots, N]$ is the sample number in the y direction[5].

Likewise, the projection data recorded by the scanner (in the form of sinograms) must be discrete in nature. Assume that these data are represented by $p_s[a, b]$, where $a \in [1, \ldots, A]$ and $b \in [1, \ldots, B]$ are the sample numbers in the θ and r directions, respectively. We can think of each $p_s[a, b]$ observation specifying a thin "strip" in the aperture along which the line integral of activity is computed, as shown in Figure 2.5.

It is important to realize that since we now have a finite number of discrete noisy projection samples, it is no longer possible to perfectly reconstruct $f(x, y)$ using equation 2.4. Because of this, we focus on the discrete inverse Radon transform operator, $\mathcal{R}_d^{-1}$, that uses the noisy projection data to produce a sampled reconstructed image $f_{rec}[m, n]$.

$$
f_{rec}[m,n] = \mathcal{R}_d^{-1}\{p_s[a, b]\}
\tag{2.6}
$$

For obvious reasons, we require the reconstruction $f_{rec}[m, n]$ to approximate, as closely

---

[5]In this thesis, we follow the well-established convention of using square brackets to denote discrete sequences (such as $f[n]$), to distinguish them from functions of a continuous variable such as $f(r)$.

as possible, the samples $f_s[m, n]$ of the true radiotracer distribution $f(x, y)$. There are two common techniques used in clinical practice to do this - filtered backprojection and iterative reconstruction, as explained below.

2.4.1 Filtered Backprojection (FBP)

This is the traditional method still commonly used due to its computational simplicity. PET images are typically characterized as having predominantly low frequency content, due to the underlying structure. The noise in PET images however, tends to have significant energy at high spatial frequencies, leading to a serious problem when applying the ramp filter $R(\omega)$ of equation 2.5, whose response increases with increasing ω [5]. The high frequency noisy components undergo a large amount of amplification. In order to reduce the effect of the noise term, $R(\omega)$ is modulated by a window function $W(\omega)$ that tapers off at high ω to form:

$$R_{fbp}(\omega) = R(\omega)W(\omega) \tag{2.7}$$

Common choices for W include windows used in other areas of digital signal processing (DSP), *e.g.* Butterworth, Hann, and Hamming. Replacing $R(\omega)$ in equation 2.5 with the windowed ramp filter $R_{fbp}(\omega)$ gives the expression used to reconstruct an image with FBP.[6]

FBP reconstruction has the desired effect of attenuating the noise arising from high ω, but it also filters out any useful signal content that might be present in this part of the spectrum. Increased spatial resolution comes at the cost of decreased noise suppression. In addition, FBP reconstructed images suffer from a radial "streaking" artifact, which results from noise in the projection data [1]. Although FBP can

---

[6]It should also be noted that there is an alternate way of performing FBP called *convolution backprojection*, in which the filtering described is performed in the *spatial domain* instead of the frequency domain.

give reasonable results (and is used in practice due to its simplicity and ease of implementation), the noise properties are far from optimal. An example of an image reconstructed with FBP is shown in Figure 2.7.

2.4.2 Iterative Reconstruction

This method of image reconstruction is more computationally intensive than FBP, but can lead to better quality images. Another key advantage of iterative reconstruction methods is that they allow more sophisticated modeling of noise than possible with FBP.

Unlike FBP, iterative reconstruction does not use an analytic expression to arrive at a reconstructed image. The technique is based on an increasingly refined series of image estimates, which may include any known *a priori* information (such as the exclusion of negative numbers in the image, which are physically meaningless for a radiotracer distribution). After each estimate is generated, it is forward-projected. The projections resulting from the estimate and the actual measured projection data are then compared, using an objective function quantifying the discrepancy. The estimate is then adjusted, and these iterations continue until the objective function has been minimized, or a suitable result has been obtained [1]. Generally, this process requires the development of:

- a model that accurately describes the (noisy) data acquisition system,

- an objective function that provides a numerical measure of "goodness of fit" of measured data to the estimate, and

- an algorithm that refines the image estimate for the next iteration.

The most popular of these techniques include maximum–likelihood expectation–maximization (MLEM) and ordered–subset expectation–maximization (OSEM). The

OSEM algorithm, originally proposed by Hudson and Larkin, is based on dividing the projection data into subsets, and provides very fast convergence times while not sacrificing image reconstruction quality. An extensive description of OSEM is beyond the scope of this thesis, but the reader is referred to the original paper for details [8]. An OSEM reconstructed cardiac PET image is shown in Figure 2.7. Note the different noise properties, such as the absence of FBP-like streaking artifacts in the OSEM image.

a   b
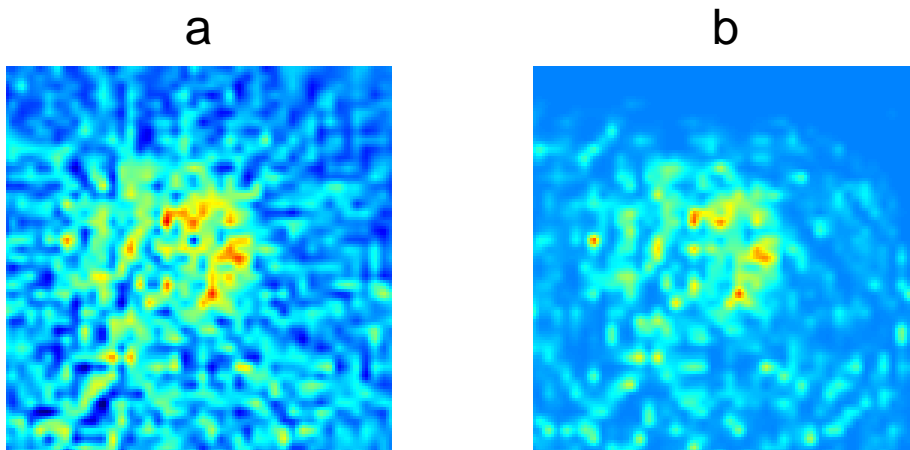
Figure 2.7: a) A noisy FBP reconstruction, and b) a noisy OSEM reconstruction of the same sinogram data, showing the very different noise characteristics.

We must emphasize that though the preceding discussion focused on 2D, this study is concerned with the entire myocardial volume. Hence, our image representation must encompass all three spatial dimensions. This is done with the use of volumetric *tomography*, a technique that represents a 3D volume as a "stack" of 2D cross-sectional images, called *slices*, taken at discrete points in the z direction. As mentioned previously, each image slice is generated independently from 2D projection data in the OHI PET scanner.

2.5   Sources of Noise in PET Images

Because this work is concerned with noise removal from PET images, we summarize in this section the most important sources of error that cause the reconstructed PET images to differ from the actual radiotracer distribution in the body. The preceding description of PET physics, scanner hardware, and image reconstruction has revealed several phenomena that degrade the images, namely:

- positron range, leading to loss of resolution

- the reconstruction software's assumption that the annihilation photons are ejected at *exactly* 180 degrees apart, leading to loss of resolution

- scattered counts, leading to counts at erroneous positions

- random counts, leading to counts at erroneous positions

- the effects of detector dead time, leading to loss of counts

- loss of photons that do not hit the detectors (out of field counts)

In addition to these, the following effects (which we will not describe in detail) introduce additional noise into the reconstructed PET images [1]:

- radial blurring (due to the use of thin, long detector crystals)

- attenuation (which leads to underestimation of activity inside the body)

- poisson counting noise

- patient and organ motion during scan time

- image reconstruction artifacts (*e.g.* FBP streaking)

Upon inspecting Figure 2.7, we see that the accumulation of all of these effects contributes such image degradations as:

- 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

The statistical properties of the noise in cardiac PET data are discussed and a model that can be used to characterize this noise is given in Chapter 5.

In the next chapter, we explain how the principles of PET imaging are applied in an application of cardiology.

Chapter 3

Background — PET in Cardiology

The previous chapter explained the PET image formation process and discussed the inherent noise that contaminates these images. This chapter provides the reader with relevant background details on how PET is used in cardiology, in order for the reader to understand the PET application considered in this thesis. Section 3.1 gives a brief description of some relevant aspects of cardiac anatomy and function. Following this in Section 3.2, we provide an explanation of the different spatiotemporal representations of the heart. Since the data of interest are time series of images, we describe dynamic PET studies in Section 3.3, and explain the need for quantitative cardiac PET studies. Finally, we provide an overview of the method by which perfusion estimates are extracted from the measured data with compartmental modeling.

3.1   Cardiac Anatomy and Function

Cardiovascular diseases are a major cause of morbidity and mortality and have an enormous personal, community and health care cost. Cardiovascular diseases (including heart disease and stroke) are the leading causes of death in Canada (34%) [9]. Further effects of these illnesses are borne by those living with the disease, through the impact on quality of life and disability. It is estimated that cardiovascular dis-

ease costs $18.4 billion dollars annually in Canada, including direct and indirect costs. Consequently, the societal and economical motivations for accurate cardiac PET studies that might help in the diagnosis and treatment of CAD are self-evident.

As shown in Figure 3.1, the human heart consists of 4 chambers - the left and right atria, and the left and right ventricles - which are separated by valves which prevent the backflow of blood between chambers. During the cardiac cycle (*i.e.* duration of one heartbeat), the atria collect blood that is returned to the heart through veins from the previous cycle. The left atrium holds oxygenated blood that returns from the lungs (via the pulmonary veins) and the right atrium holds deoxygenated blood being returned from the rest of the body. The ventricles are responsible for pumping blood to the body through a system of arteries. The right ventricle sends blood to the lungs via the pulmonary artery. The left ventricle (LV) is responsible for pumping oxygenated blood to the whole body via the aorta, which then repeatedly branches into successively smaller arterial blood vessels. This pumping is performed by the contraction of the muscular wall surrounding the LV (the *myocardium*). Hence, the ability for the LV to forcefully contract is of vital importance, as it provides the blood supply for the entire body. For this reason, cardiac PET studies seek to measure the degree of LV myocardial perfusion as a measure of the ability of the muscle to perform this critical role [10].

Typically, cardiac PET images are taken under two conditions – *rest* and *stress*[7]. High levels of perfusion to specific myocardial regions confirm that the muscle is *viable*, or able to perform its role in that area. A lower level of perfusion indicates a problem with blood supply to that part of the myocardium. If the same area of restricted perfusion is detected during both rest and stress conditions, a permanent infarction has most likely occurred, resulting in scar tissue. However, if the restriction

---

[7]Normally, a drug such as *dipyridamole* is administered to simulate exercise conditions (stress), in order to avoid practical complications that might affect image quality (*e.g.* excessive patient motion and short tracer half-life).
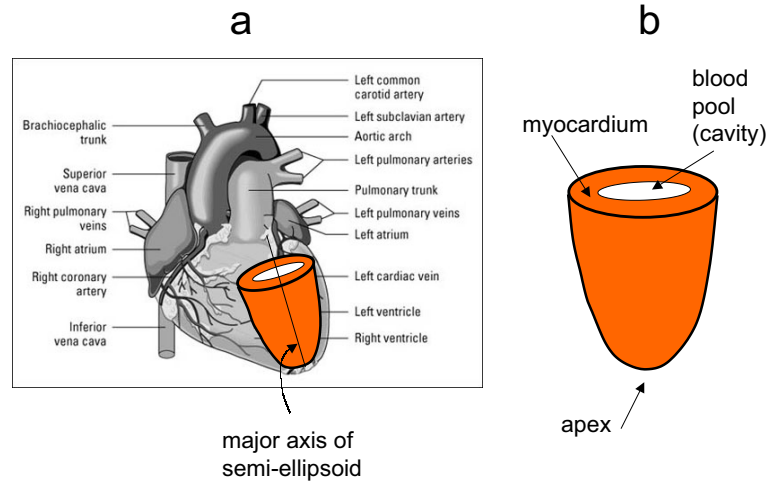
Figure 3.1: a) Diagram of human heart with relevant structures iden-
tified, b) Diagram of left ventricle with relevant structures
identified, showing its orientation with respect to the body
and PET scanner axes

only appears under stress (but not rest), then the cause is likely *reversible ischaemia*,
a condition that is treatable with surgery [2]. In general, PET images that are taken
under stress conditions are better quality than those taken under rest conditions.
This is because of the higher myocardial perfusion and tracer accumulation in a
stress study, thereby increasing the signal to noise ratio of the raw image.

The tissue being considered in this study will be of the LV and surrounding area.
As seen in Figure 3.1, the LV consists of a muscular wall which surrounds a *blood
pool*, or *cavity*. The shape of the myocardium is modeled as a semi–ellipsoid, with
major axis passing through the apex and the center of the ventricular base.

3.2   Representations of the Heart used in PET

There are two common image orientations of the heart muscle, depending on the

manner in which the tomographic slices of the myocardium are generated.

- **Short axis (SA) view**

  In this case, the image slices are formed from the intersection of a thinly spaced series of planes perpendicular to the major axis of the LV's modeled ellipsoid. For this reason, the myocardium tissue appears as an annular region in each image, with the diameter of the inner and outer rings decreasing for those slices that approach the apex. This is shown in Figure 3.2.
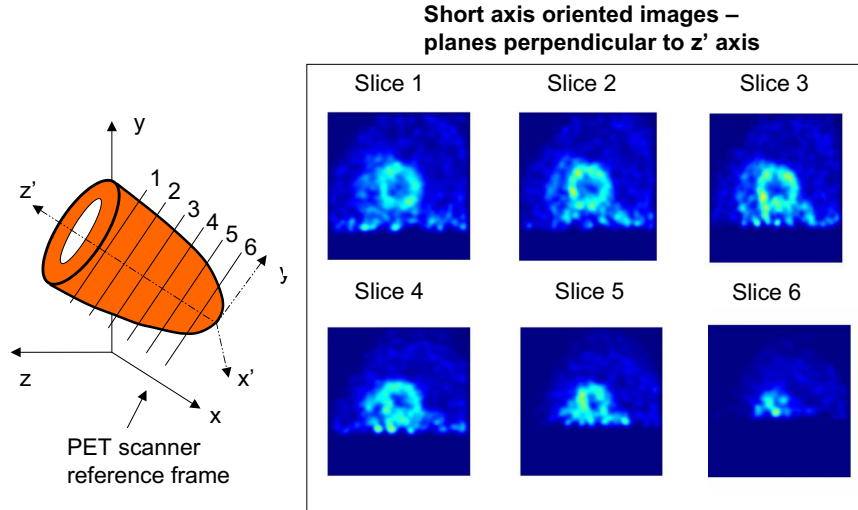
Figure 3.2: Image slices of left ventricle in short axis orientations

- **Transaxial view**

  In this case, the image slices are formed from the intersection of a thinly spaced series of planes perpendicular to the z axis of the PET scanner. The LV's position in the chest is "tilted" with respect to the z axis of the PET scanner, so the myocardium appears as a "horseshoe" shape in each image due to the the oblique angle that the myocardium forms with this axis.

- **Cardiac Polar Map**

    The *cardiac polar map* is an alternate representation commonly used by cardi-
    ologists. As shown in Figure 3.3, this is a projection of the entire myocardium
    volume onto a two-dimensional image. The viewpoint is along the major axis
    of the ellipsoid model, *as seen from the apex.* In this representation, a circum-
    ferential profile from each SA-oriented slice is represented as a thin ring in the
    polar map. The concentric rings in the middle of the polar map correspond
    to the slices near the apex, while the outer rings represent those around the
    ventricular base. Each ring is divided into *sectors* around its circumference,
    and the colour value in each sector denotes the sum of radiotracer activity in
    the corresponding region for that specific slice.
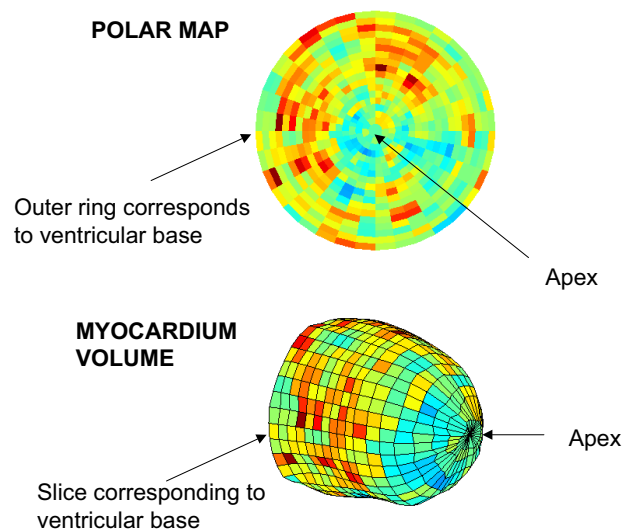
Figure 3.3: Cardiac polar map and associated viewpoint

3.3 Quantitative Dynamic PET Studies

Applying the image reconstruction techniques described in Chapter 2 will give an image of the three–dimensional spatial distribution of radiotracer present in the patient's body. This is formed from an accumulation of photon counts at the scanner's detector array over a *fixed time*. Such a static representation can be useful to a cardiologist, and is sufficient to perform a *qualitative* analysis. For example, the image in Figure 3.4 clearly shows the approximate size, shape and location of a myocardial defect. However, this qualitative information is not sophisticated enough to use in a physiological model that extracts *quantitative* diagnostic parameters, such as myocardial perfusion.
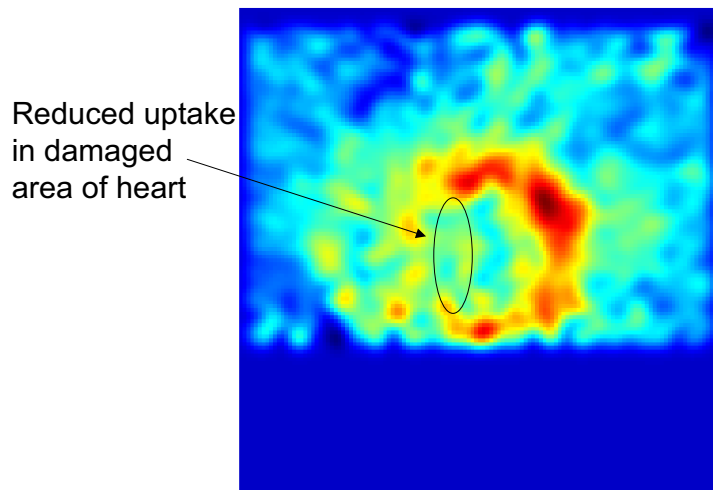
Figure 3.4: Short axis oriented PET image of infarcted tissue

Dynamic PET studies are concerned with how the tracer distribution changes with time, giving the information required to measure the *in vivo* physiological process of interest. A time-varying representation of the radiotracer activity requires several images taken in succession, the reconstruction of which is independent from the others.

Each 3D image has associated with it a time interval (called a *frame*) during which the counts occurring during that frame are used to generate that image only. During the study, the tracer injected into the bloodstream will be carried by a variety of transport mechanisms to the different regions around the heart. For a ^{82}Rb study (which lasts 10 min), there are 17 frames, structured as shown in Table 3.1.

Frame	*Frame Duration*	*Time (start of study is $t = 0$)*
1-12	10 seconds	from $t = 0s$ to $t = 120s$
13-14	30 seconds	from $t = 120s$ to $t = 180s$
15	60 seconds	from $t = 180s$ to $t = 240s$
16	120 seconds	from $t = 240s$ to $t = 360s$
17	240 seconds	from $t = 360s$ to $t = 600s$

Table 3.1: Frame times and durations for dynamic ^{82}Rb PET study

Figure 3.5 is an image sequence showing the variation in radiotracer uptake over time for a midventricular transaxial-oriented slice. Notice that as the study progresses, there is transfer of ^{82}Rb activity from the heart chambers to the horseshoe–shaped myocardium region.

There are two alternate representations of PET data relevant to dynamic studies:

- **Time activity curves (TAC)**

  A TAC is a way of showing the radiotracer activity present in a specific anatomical area (called a *region of interest* (ROI)) over time. To do this, the myocardium is segmented into different ROIs, which are then superimposed onto each image in the time series. Summing the pixel activity within an ROI (for a given time frame), and then repeating for all time frames, generates a one dimensional signal (a TAC) that shows the radioactivity of the ROI as a function of time for the duration of the study. Depending on the location of the ROI selected, TACs can be generated for different parts of the heart. In this thesis,
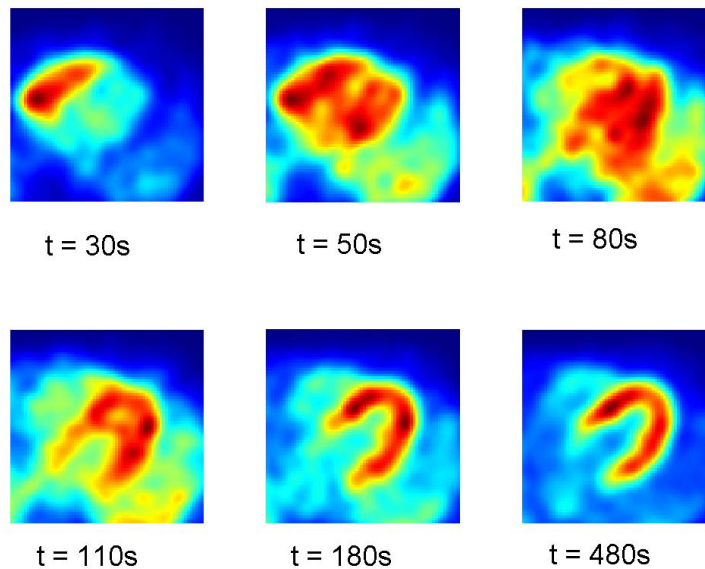
Figure 3.5: Image sequence showing increased myocardial uptake as study progresses (transaxial view). Frame times are indicated below each image.

we deal with tissue TACs (in which the ROIs are located in the myocardium) and blood TACs (in which the ROIs are located in the LV blood pool). Typical TACs (for tissue and blood) are shown in Figure 3.6. Note that the reason for the faster sampling at the beginning of the study is to capture the TAC uptake transient (with significant high frequency components) that occurs close to $t = 0$ s. It is evident that as the study progresses, activity is transferred from the blood to the myocardium.

- **Dynamic Polar Map**

  A dynamic polar map is a time-indexed set of cardiac polar map images, in which each polar map gives the sum of activity during a specific time frame. Each polar map sector (or collection of adjacent sectors) can define a TAC for that portion of the myocardium. The sectors are considered as ROIs placed circumferentially around the myocardium on SA images
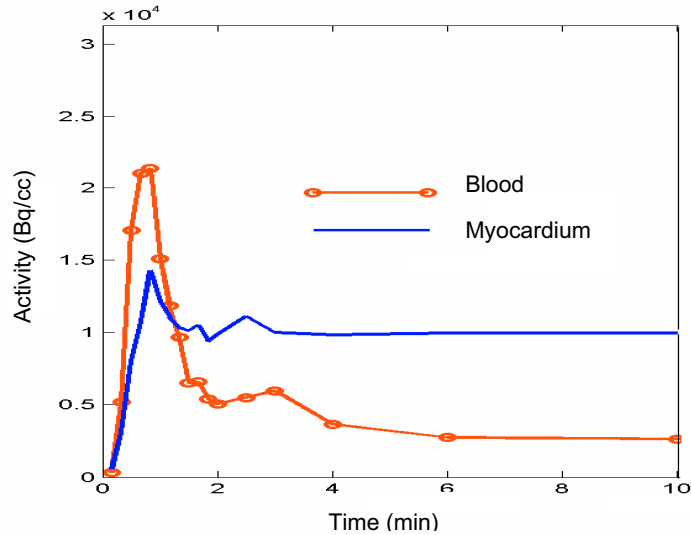
Figure 3.6: An example of a time activity curve, showing myocardial uptake during a dynamic PET study.

3.4 Tracer Kinetics and Compartmental Modeling

In order to arrive at a quantitative measure of myocardial perfusion (measured in mL/min per gram of tissue), it is necessary to use a model relating the spatiotemporal distribution of radiotracer in the body (measured by the PET scanner) to the flow parameter of interest. Such a model is called a *tracer kinetic model* and incorporates the underlying physiological processes in the body.

When the radiotracer is injected into a patient's bloodstream, this "labeled" blood returns to the LV cavity (after being oxygenated) via the circulatory system, as described at the beginning of this chapter. Subsequent contraction of the LV forces the labeled blood throughout the body via arteries – among these are coronary arteries that supply blood to the LV myocardium tissue. These eventually branch off into thousands of capillaries present in the myocardial wall. In order for the blood to

travel from these capillaries to the myocardial muscle cells, they have to travel across the capillary and cell membranes.
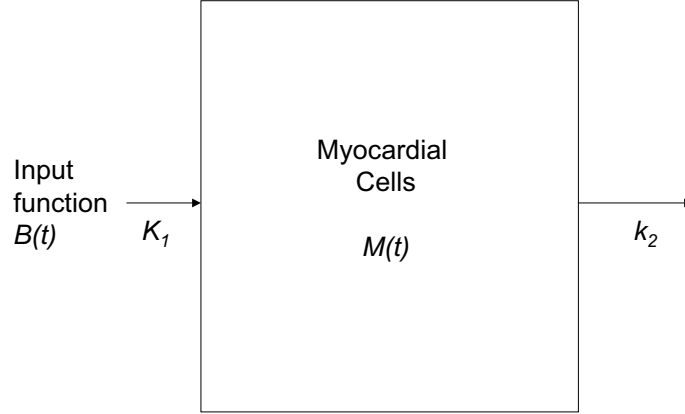
Figure 3.7: The one-compartment model used to quantify myocardial perfusion from a dynamic ^{82}Rb PET study.

The above description forms the basis of various compartmental models that can be used to model this dynamic system [11]. In this work, we use a simplified *one-compartment* model described in [4] and shown in Figure 3.7. The input function to the model, $B(t)$, corresponds to the time activity curve for the blood pool region (since this is the *source* of the myocardial uptake). This model does not consider separate capillary and interstitial spaces, instead using a single compartment to represent myocardial tissue. We denote the time-varying uptake of blood in this compartment as $M(t)$. The exchange of blood to and from this compartment is governed by the following rate constant parameters:

- K_1 (**Uptake**) - this characterizes the transfer of blood from the LV blood pool to the myocardium ($mL/min/g$)

- k_2 (**Washout**) - this characterizes the evacuation of excess blood from the

myocardium (min^{-1}).

The required mass balance results in the following differential equation:

$$\frac{d}{dt}M(t) = K_1 B(t) - k_2 M(t) \tag{3.1}$$

which can be solved for $M(t)$:

$$M(t) = K_1 \int_0^t e^{-k_2\tau} B(t-\tau)d\tau = K_1 e^{-k_2 t} * B(t) \tag{3.2}$$

By segmenting the measured PET data (f) into myocardium and blood pool ROIs, $B(t)$ and $M(t)$ can be estimated from the TACs in these areas. Of course, these data do not give instantaneous concentration values, but rather an integration of counts over time frames. Also, the measured PET data in the myocardial ROI consist of counts due to both *a)* tracer taken up by the myocardium muscle tissue and *b)* tracer in blood within the myocardium (*i.e.* in the arteries). This is accounted for by introducing the TBV (total blood volume) parameter and incorporated into the above model as follows:

$$f(t) = (1 - TBV) \cdot M(t) + TBV \cdot B(t) \tag{3.3}$$

where $f(t)$ refers to the measured PET data at the polar map sector being processed.

The one-compartment model describes a hypothesized "ideal" relationship between the rates of change of tracer concentrations $B(t)$ and $M(t)$, which is governed in equations 3.2 and 3.3 by the quantities K_1, K_2 and TBV. It is important to consider that these parameters vary throughout the myocardium, so they are computed at *each sector* in the cardiac polar map. They are determined with a least-squares analysis, as described below.

The parameters are given arbitrary starting values. With these, the TACs result-

ing from the model's equation (the right hand side of equation 3.3) are compared to the TACs for the measured PET data (the left hand side of equation 3.3). An residual error (measuring the goodness of fit between the model and the measurement) is calculated and recorded. The constants are then modified, and the new model's residuals are determined. This process is repeated until the sum of squares of the calculated residual represents a minimum, indicating good agreement between the model and PET measurements. To ensure fast convergence, the adjustment of the rate constants is a sophisticated gradient-based process (*e.g.* using the "large-scale: trust-region reflective Newton" algorithm in the Matlab *lsqnonlin* function [12]). Figure 3.8 shows an example output of the one-compartment ^{82}Rb quantitative model used at OHI, showing how the parameters of interest vary throughout the myocardium.
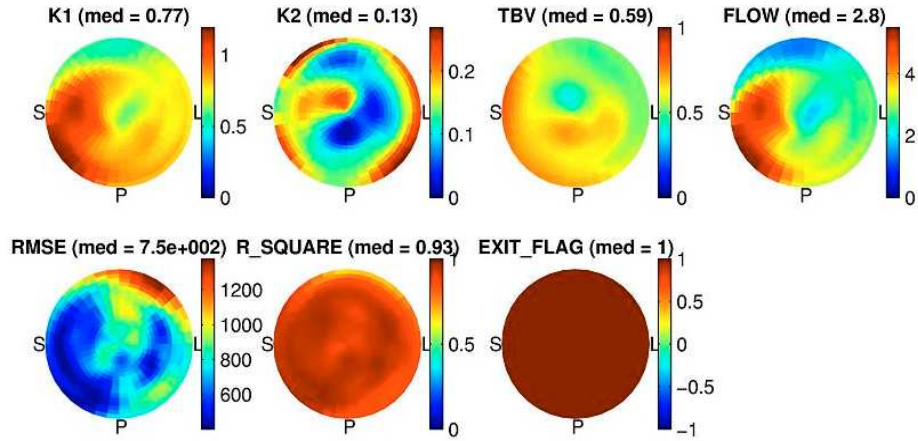
Figure 3.8: Example of one-compartment ^{82}Rb quantitative model output.

In the next chapter, we focus on the problem of image noise reduction and explain how wavelet–based strategies can be used to advantage for cardiac PET data.

Chapter 4

Image Denoising

The previous two chapters have given the reader an overview of PET and its use in cardiology, and described various sources of noise in PET data. In this chapter, we turn our attention to the techniques that can suppress this noise. Section 4.1 gives a review of classical noise reduction methods. This is followed in Section 4.2 by a brief discussion of why these methods are not optimal for PET images, and how these shortcomings can be overcome by using wavelet-based methods. In Section 4.3, a detailed review of the wavelet denoising literature is presented. We conclude this chapter by providing a justification and details of the specific wavelet representation used in this thesis – the spline-based *discrete dyadic wavelet transform*.

4.1 Classical Techniques for Image Denoising

Image denoising (also known as *image noise suppression* or *image restoration*) is a well-studied area.[8] Accordingly, several techniques have been proposed for eliminating

---

[8]We distinguish image denoising from *image enhancement*, though these are sometimes used in the literature to describe closely related techniques. Generally, image enhancement is used to make an output image that is more *subjectively pleasing* to the observer using, for example, histogram equalization (to increase contrast in an otherwise dull image) or unsharp masking. Image denoising, however, generally uses known properties of the noise degradation to recover the original image and can usually be measured objectively with a numerical figure of merit.

noise from images. These have their own merits and downfalls, depending on both the type of image being denoised and the characteristics of the noise degrading the images. After discussing a suitable noise model, we present a brief overview of some well-known noise filters in this section, summarized primarily from [13].

4.1.1   Noise Model and Properties

The most common noise model is the additive noise model, in which we observe f_{noisy}, the combination of a "true" underlying image f (that we are trying to recover) and a noise contamination term n (that we are trying to remove):

$$f_{noisy}[k] = f[k] + n[k] \tag{4.1}$$

We use a single lexographic index $k \in \{1, 2, ..., N\}$ that varies over all N "pixels" in the representation, in order that the noise model may be applied in general to multi-dimensional signals. In each case, the value of n at each k value is a random variable that follows a probability density function (PDF), for which important characterizing values are its mean and variance.

The N x N covariance matrix of n is:

$$\mathbf{C_n} = \begin{pmatrix} \sigma_{1,1} & \cdots & \sigma_{1,N} \\ \vdots & \ddots & \vdots \\ \sigma_{N,1} & \cdots & \sigma_{N,N} \end{pmatrix} \tag{4.2}$$

with each entry the covariance between pairs of random variables (*i.e.* samples of n):

$$\sigma_{i,j} = E[(n_i - E(n_i))(n_j - E(n_j))] \tag{4.3}$$

where the E operator represents expectation. If $\mathbf{C_n}$ is a diagonal matrix, we have

uncorrelated noise, the diagonal elements being the individual variances of the noise samples. If $\mathbf{C_n}$ has non-zero off-diagonal elements, the noise is said to be *correlated*, indicating that the noise values depend to some degree on the position in the image. The $\mathbf{C_n}$ matrix describes the *texture* of the noise - a "sparse" matrix (one with isolated off-diagonal elements) will give a more grainy looking noise contribution than that resulting from a more dense version, which would appear as larger "blobs" or streaky artifacts in the image.

A special case of interest is when the noise samples are independently and identically distributed (i.i.d.) with a zero mean Gaussian distribution, *i.e.* they follow the PDF:

$$p_n(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-x^2/2\sigma^2} \tag{4.4}$$

where σ^2 is the variance of the noise.

The goal of image denoising is to obtain $f_{denoised}$, an estimate of f, based on the noisy observation f_{noisy} and any known properties of n (*e.g.* mean, variance, PDF).

4.1.2 Low Pass Filtering

A simple yet effective method of noise reduction can be done with a *smoothing filter*. This is an intuitive approach that reduces the effect of noisy pixel values by introducing controlled blurring. Smoothing filters can be designed and applied in the spatial domain by performing spatial convolution of the noisy image with an appropriately selected filter kernel. For each pixel in the denoised image, the value of $f_{denoised}$ is a function of the other pixel values in the kernel. Alternatively, smoothing filters may also be designed and applied in the frequency domain (since image smoothing is a low-pass operation) with the use of two-dimensional Fourier transform techniques. Regardless of its implementation, the most important factor affecting performance is the amount of smoothing applied. Insufficient smoothing leads to inadequate noise

suppression, while oversmoothing leads to loss of important edge information. This is because both noisy areas and edge areas consist of sharp transitions in pixel values, and both are highly affected by the smoothing operation. In order to successfully use this type of filter, the inherent tradeoff between noise reduction and loss of detail must be acceptable for the given application.

A smoothing filter of particular importance in this thesis is the Gaussian lowpass filter (GLPF), which has the following impulse response in 2D:

$$h(x, y) = k_0 e^{-(x^2+y^2)/2\sigma^2} \tag{4.5}$$

and 3D:

$$h(x, y, z) = k_0 e^{-(x^2+y^2+z^2)/2\sigma^2} \tag{4.6}$$
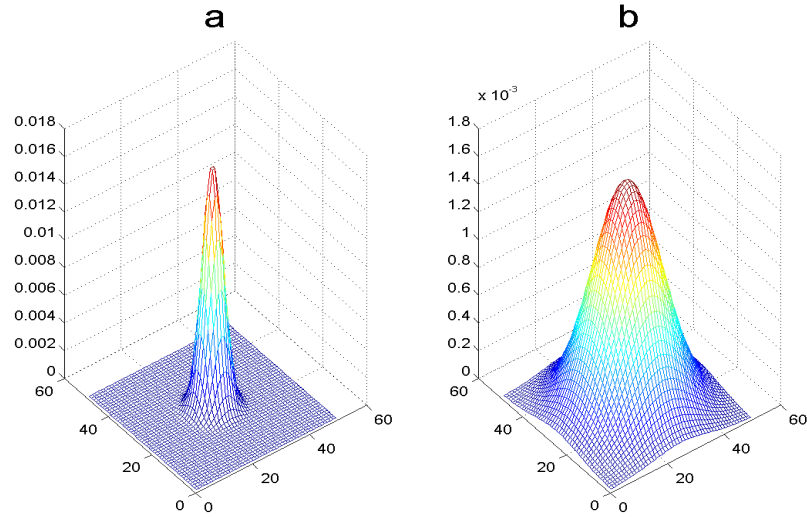
Figure 4.1: Impulse response of a 2D GLPF with a) $\sigma = 3$, b) $\sigma = 10$. The k_0 constant has been selected such that the total "volume" under the curve is equal to one.

The σ parameter controls the width of the Gaussian kernel (see Figure 4.1), with

higher values of σ leading to increased smoothing in the denoised image. This is illustrated in Figure 4.2 on the 512x512 *Lena* test image.

Figure 4.2: Filtered version of noisy "Lena" image using the filters in Figure 4.1 - a) original image, b) noisy image with $\sigma_n = 0.005$, c) GLPF with $\sigma = 3$ applied, d) GLPF with $\sigma = 10$ applied

There are several advantages to using a GLPF:

- it can be reformulated into a *separable* configuration, since the above impulse responses can be rewritten as products of one dimensional functions. This allows very simple and efficient denoising by successively filtering the data along the row, column (and slice) directions

- the GLPF is a real-valued function in both space and frequency, allowing simplified processing by avoiding the need for complex numbers

- if a frequency domain design and/or implementation are desired, the GLPF has the unique property of having the same basic bell-curve shape in both domains. This behaviour makes it intuitive when investigating the effect of changing the σ parameter - as σ increases in one domain (widening the bell curve), it decreases in the other.

Smoothing filters in general are effective at removing several noise types, most notably Gaussian and uniform noise[13].

4.1.3 Order Statistic Filters

Other classical noise reduction methods are based on order statistics of the image. Unlike a smoothing filter, the value of the pixel being processed depends on the *sorting order* of adjacent pixels. The simplest (yet quite effective) example is the *median filter*, which denoises a pixel by replacing it with the median value of the pixel values in the filter neighbourhood $\mathcal{N}$:

$$f_{denoised}[k] = \underset{a \in \mathcal{N}_k}{median}\{f_{noisy}[a]\} \tag{4.7}$$

The median filter can result in better quality images (less blurring) than those denoised with smoothing filters for certain noise types. It is especially effective for *impulse* noise or *salt and pepper* noise, characterized by isolated extreme values in the image. Adaptive approaches to median filtering react to changing image characteristics by varying the size of the neighbourhood used. As a testimony to its effectiveness, more sophisticated refinements on the basic median filter are still appearing in the literature, for example [14, 15].

4.1.4   Wiener Filters

This image denoising technique (also known as the minimum mean-squared error (MMSE) filter) uses a slightly different model of noise contamination than the one above. The ideal image f is assumed to have been filtered by a linear, shift-invariant degradation function h before the addition of noise n, resulting in the observed signal f_{noisy}:

$$f_{noisy} = f * h + n \tag{4.8}$$

Both the image f and noise n are assumed to be uncorrelated random processes, and one of them is assumed to have zero mean. In order to get a $f_{denoised}$ estimate as close as possible to f, the filter attempts to minimize the following error metric:

$$e^2 = E[f - f_{denoised}]^2 \tag{4.9}$$

The image estimate that does this is given in the Fourier domain as:

$$F_{denoised} = \left(\frac{1}{H} \frac{|H|^2}{|H|^2 + S_n/S_f} \right) F_{noisy} \tag{4.10}$$

where the uppercase quantities are the Fourier transforms of those in equation 4.8, $S_n = |N|^2$ is the power spectrum of the noise, and $S_f = |F|^2$ is the power spectrum of the noisefree image. Because the S_n and S_f quantities are unknown in many applications, their ratio is often assumed to be a constant, and this value is chosen manually to give the best subjective denoised result. Wiener filtering is considerably more complicated than the other methods described above, and the suitability of this method depends on the validity of the assumptions made[13].

4.2 Wavelet-based Image Denoising

4.2.1 Motivation

A common image noise reduction method used at OHI involves spatial convolution of the volume of data with a fixed width, three-dimensional Gaussian kernel. It produces suitable results in many cases, but does not provide optimal separation of signal from noise. This is because the amount of smoothing is constant over the *entire image*, which consists of a combination of slowly-varying areas interspersed with important discontinuities. Since there is no way for this method to *adapt* to changing image characteristics, increased noise suppression comes at the cost of loss of detail. To achieve satisfactory noise reduction, we lose important small-scale image features, a situation that is undesirable when extracting precise regional flow measurements from these data, as is the case in myocardial perfusion studies.

An example is shown in Figure 4.3, where a portion of the standard *Lena* test image has been corrupted with additive Gaussian noise. In order to suppress the noise that is prominent in the shoulder area (a smooth part of the image), we need to increase the width of the Gaussian filter kernel. Though effective at removing noise in this region, we notice that this operation causes undesired blurring effects of the smaller-scale features in other more detailed areas. For example, the pupil of the eye cannot be discerned at this level of smoothing. This clearly shows the tradeoff between increased noise suppression and the resulting loss of detail.

As mentioned in the introduction, the signals that arise in nuclear medicine (such as PET images and TACs) are not ideally suited to analysis with traditional frequency domain (Fourier) techniques. This is because, in general, such signals are *non-stationary* and therefore do not exhibit global, periodic behaviour. Instead, relevant signal features are localized in both space and frequency. The basis functions of Fourier analysis are sinusoids of a given frequency, defined over the entire signal
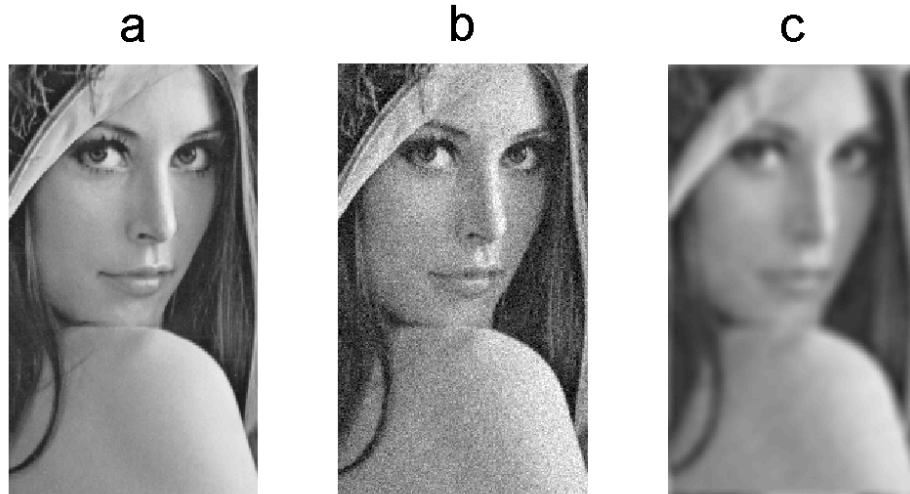
Figure 4.3: Lena image with fixed width Gaussian kernel filtering – a)
noise free image, b) noisy image, and c) smoothed image

domain. Put another way, they have perfect localization in the frequency domain, but no localization in the space domain. For this reason, it is difficult to efficiently represent transient features (such as cardiac tissue defects in an image, or an uptake spike in a TAC) with such a basis. We require an alternate signal representation which allows *joint resolution* in both spatial and frequency domains.

The wavelet transform is a tool that possesses such a capability, because the basis functions are of finite duration and varying *scale* (a parameter related to frequency, as explained below). This can be likened to the information in a musical score, where information about both frequency (which note is played) and time (when and how long a note is played for) are present. Wavelets comprise one means of performing a *multiresolution analysis* (MRA) - an examination of signals at differing resolutions. This is significant, since having such information makes it possible to process features of varying size. Wavelet theory is now very well developed, and applications of this technique (especially in the fields of image compression and denoising) have been

investigated extensively in the literature during the last decade. Wavelet methods have been successfully applied in a wide range of biomedical applications, an excellent overview of which appears in [16]. A more recent review focuses on the success of using wavelet techniques in medical imaging [17]. In addition, since wavelet analysis can be applied to multidimensional signals, it is well suited to this study.

4.2.2   Wavelet Transform

This section is a very brief discussion of wavelet concepts relevant to the image denoising application. A more thorough tutorial of basic wavelet theory is provided in Appendix A, and a detailed description of the wavelet representation used in this thesis (the DDWT) appears in Section 4.4.

In a wavelet representation, a *mother wavelet* $\psi(x)$ is used to form the basis functions for the signal expansion. These basis functions are related to one another through two parameters - *scale*, s, and *translation*, τ. The wavelet transform measures the degree to which an input signal is correlated with these scaled and translated basis functions. In the case of a discrete wavelet transform (DWT), the s and τ parameters are sampled along discrete sequences. Therefore, the DWT of an input signal consists of discrete sets of coefficients. The use of a scaling function $\phi(x)$ is required in order to put a practical lower limit on the scales used in the representation. Because the DWT is most often implemented as an iterated structure consisting of cascaded low-pass and high-pass filters (using the fast wavelet transform (FWT)), the *analysis* section of the filter bank structure groups DWT coefficients into the following:

- **Approximation coefficient subbands**

  These subbands of coefficients, calculated at each level of the filter bank structure, provide a low resolution approximation of the signal being represented. As the level number increases, this approximation becomes increasingly coarse.

- **Detail coefficient subbands**

  These subbands of coefficients, calculated at each level of the filter bank structure, provides details corresponding to the difference in information between the approximations at adjacent levels. As the number of levels increases, these details become increasingly coarse.

Having generated these coefficients (which represent the input signal in the wavelet basis), it is desirable to modify them in some way, based on the application at hand. In image denoising, for example, the noise energy in the input signal is generally spread over an entire subband of coefficients, so a thresholding operation is applied in order to suppress this portion. Once the DWT coefficients have been changed, it is necessary to reconstruct the output signal. This is done by applying the inverse discrete wavelet transform (IDWT), also implemented with an iterated filter bank (in this case, the *synthesis* section).

Figure 4.4 shows a general DWT implemented with a 3-level filter bank, showing both analysis and synthesis sections.

4.3 Wavelet-based Noise Removal

Most active among denoising research efforts that seek to preserve image details over multiple scales is *wavelet-based denoising*. Techniques based on a wavelet representation exploit the wavelet's natural ability to separate signal from noise over different image scales. This provides an inherent advantage when dealing with non-stationary signals – the inclusion of localized "fine scale" functions in the basis allows one to better discern diagnostically significant details.

A thorough overview of wavelet denoising appears in [18]. The seminal work in this field was performed by Donoho and Johnstone [19, 20, 21]. They proposed thresholding of the detail coefficients of an orthogonal DWT (denoted with the $\mathcal{W}_{DWT}$
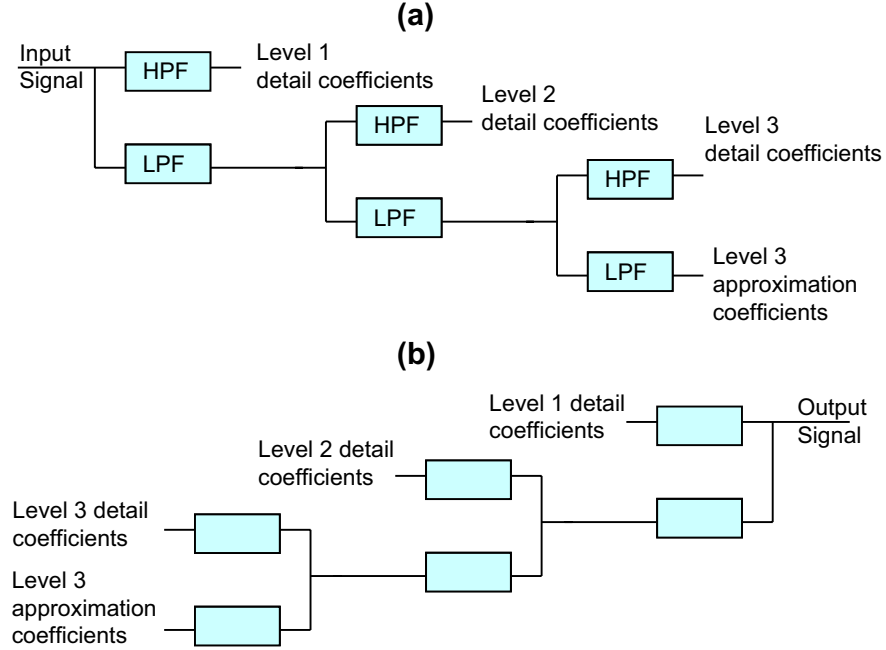
Figure 4.4: General filter bank showing wavelet coefficients - a) Analysis
filter bank and b) Synthesis filter bank. LPF = low pass
filter, HPF = high pass filter

operator) of a noisy signal f_{noisy}. This simple method is motivated by the DWT's
ability to compact most of the signal energy in a small number of high amplitude
coefficients, which are retained. The noise contribution, assumed to be due to lower
amplitude coefficients, are spread out over each detail subband and suppressed during
the thresholding step. The threshold value that is commonly used is the *universal
threshold*, $\sigma_n\sqrt{2\log N}$, where σ_n is the noise standard deviation (the noise is assumed
to be i.i.d. Gaussian with zero mean and variance σ_n^2), and N is the number of signal
samples. Note that other possible threshold choices are possible, as detailed in [18].
When σ_n is not explicitly available (as is often the case), a robust median estimator
developed by Donoho and Johnstone [19] can be used:

$$\widehat{\sigma}_n = median(abs(W[i]))/0.6745 \tag{4.11}$$

where i is an index over all detail coefficients W at the finest decomposition level.

This selection is based on the observation that at this most detailed level, the wavelet coefficients are primarily due to noise. Because there are very few coefficients due to signal, using the median eliminates any possible bias these might introduce. The 0.6745 constant stems from the fact that for a set of independent Gaussian random variables $\{n_i\}, i \in \{1, 2, ..., N\}$ with zero-mean and variance σ_n^2, $E[\text{Median}(\text{abs}(n_i))] \approx 0.6745\sigma$.

The most common ways of applying the threshold value to a DWT detail coefficient $W_D[k]$ are the *soft* and *hard* thresholding functions, defined as follows and shown in Figure 4.5:

- **Soft thresholding** (defined by the $\mathcal{T}_\eta^{soft}$ operator, where η is the threshold value)

$$
\mathcal{T}_\eta^{soft}\{W_D[k]\} = \begin{cases} 0 & \text{if} \quad |W_D[k]| < \eta \\ W_D[k] - \eta & \text{if} \quad W_D[k] \geq \eta \\ W_D[k] + \eta & \text{if} \quad W_D[k] \leq -\eta \end{cases}
$$

- **Hard thresholding** (defined by $\mathcal{T}_\eta^{hard}$, where η is the threshold value)

$$
\mathcal{T}_\eta^{hard}\{W_D[k]\} = \begin{cases} 0 & \text{if} \quad |W_D[k]| \leq \eta \\ W_D[k] & \text{if} \quad |W_D[k]| > \eta \end{cases}
$$

The denoised image, $f_{denoised}$, is formed from its modified wavelet coefficients with the inverse DWT (defined with the $\mathcal{W}_{DWT}^{-1}$ operator), so the overall operation (using soft thresholding) is:

$$
f_{denoised} = \mathcal{W}_{DWT}^{-1} \left\{ \mathcal{T}_\eta^{soft} \left\{ \mathcal{W}_{DWT} \left\{ f_{noisy} \right\} \right\} \right\} \tag{4.12}
$$

Soft thresholding (also known as *wavelet shrinkage*) generally performs better than hard thresholding, due to the latter's jump discontinuity at the threshold value. In this approach and its simplest variations, the threshold value is either uniform over
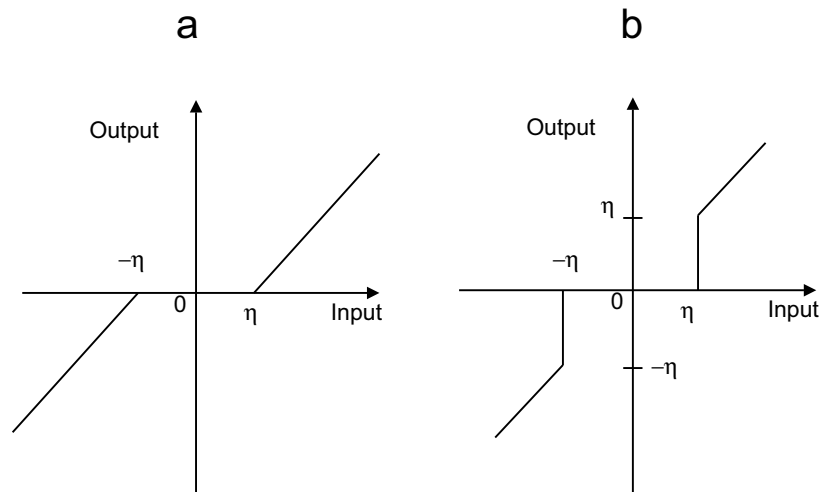
Figure 4.5: a) Soft and b) hard thresholding operators

all detail coefficient subbands at all scales, or uniform within each detail coefficient subband. The choice of which subbands to threshold and the threshold value itself varies depending on the application. Subband-adaptive constant thresholds are appropriate when the noise in the image is known to deviate from the assumption of an i.i.d. Gaussian distribution (called *correlated*, or *coloured* noise)[22]. Note that the approximation coefficients at the highest decomposition level contain important low frequency information and normally do not undergo thresholding.

Though effective in some cases, basic wavelet shrinkage of an orthogonal transform can lead to ringing artifacts in the denoised image. Coifman and Donoho later showed that a substantial improvement on this technique can be achieved if the DWT employed has the property of *translation invariance* (TI)[23]. One way of achieving this is to average the outputs that result from denoising the input image over all possible translations (called *cycle spinning*). A more computationally efficient approach involves omitting the downsampling step in the FWT filter bank implementation. Of

course, this leads to an overcomplete (redundant) wavelet representation, in which *each subband* of detail and approximation coefficients has the same number of samples as the input signal[9].

In an effort to further separate noise from signal over spatially changing image characteristics, a recent study by Chang *et al.* focused attention on the choice of threshold [24]. Rather than using either a globally constant or subband–dependent constant threshold, they show that superior denoising results can be achieved if the threshold changes *within a subband* of detail coefficients - an idea known as *spatially adaptive thresholding*. The threshold for the coefficient $W_D[k]$ in a given detail subband is :

$$\eta[k] = \frac{\sigma_n^2}{\sigma_w[k]} \tag{4.13}$$

where σ_n^2 is the noise variance in the noisy image, and $\sigma_w[k]$ is the standard deviation of the detail coefficient at position k. Each DWT coefficient $W_D[k]$ is then modified (using soft thresholding) with a threshold value *unique to that coefficient location*, $\eta[k]$. The σ_w parameter is estimated from the neighboring coefficients in a sophisticated process known as *context modeling*. This neighbourhood includes coefficients in both the subband being thresholded, and its *parent subband* (*i.e.* the coefficients at the same spatial location and subband orientation but at the next coarser scale). In areas where the coefficient variance is large, the image exhibits details (such as an edge) and the threshold resulting from equation 4.13 is lower, thereby preserving these details. Conversely, areas with smaller σ_w correspond to smoother regions and the denoising can proceed more aggressively by increasing the threshold.

This idea has been extended by Marpe *et al.*, who in addition to using a spatially adaptive threshold (albeit with a simpler context model), perform multiple iterations of the thresholding process on the wavelet coefficients [25]. Since the initial estimate of $\sigma_w[k]$ is based on noisy coefficients, it will be somewhat inaccurate. Subsequent

---

[9]See Appendix A for more details on overcomplete representations.

$\sigma_w[k]$ estimates (and the resulting adaptive thresholds based on these) are determined using denoised coefficients, which are found to be more reliable.

Most recent work by Jin *et al.* uses a refinement called *cross-scale regularization* (CSR) which increases noise suppression, particularly at the finest scale of decomposition [26, 27]. Typically, these subbands contain overwhelming amounts of noise (indeed, this is why these subbands are used for noise level estimation). Relevant signal details are present but deeply hidden in this subband, yet it is possible to "find" these relevant coefficients using information extracted from the next higher scale wavelet coefficient subband (which is not as noise contaminated). An edge map on this subband can identify the location of important image features in the finer scale subband that should be retained. The result provides information that can be used to refine the reconstruction, by including the coefficients at the same edge positions in the finer scale subband, and reducing the other coefficients, which are considered to be due to noise.

Because the data from a dynamic cardiac PET study are four-dimensional in nature, they can benefit from applying the techniques described above using higher dimensional wavelet transforms. In theory, this can lead to better denoising results by exploiting the high correlation of pixels within an image, between adjacent image slices, and between consecutive time frames. Multidimensional transforms have been used with success in other denoising applications, such as transmission electron microscopy [28] and cardiac ultrasound [29].

Turning to the image class of interest ([82]Rb cardiac PET studies), wavelet-based denoising has been used recently by Lin *et al.* with promising results [3]. A series of wavelet protocols (that are applied to uptake images, polar maps, TACs and various combinations of these) were developed and compared in this work. Each protocol used a basic, modified wavelet shrinkage technique to denoise the 1D and 2D signals. The application under investigation in this work was similar to ours – the effect of

denoising on the accuracy of perfusion measures.

In addition to denoising in the image domain, wavelets have been investigated in other areas of tomography as well. Some groups have shown that advantages can be gained from incorporating wavelet processing earlier in the reconstruction. Kolacyk [30] does this by performing a shrinkage operation on coefficients resulting from a wavelet-vaguelette decomposition introduced by Donoho [31]. Bhatia *et al.* represent the standard FBP ramp filter in a wavelet basis, thus allowing a multiscale reconstruction [32]. Sahiner and Yagle first filter the projection data (treating them as a series of 1D signals in the radial variable) with the wavelet transform and short time Fourier transform (STFT) before image reconstruction, then compare these results with those resulting from image domain denoising. It is interesting to note that they also extended this idea using iterative thresholding [33, 34]. More recently, Kalifa *et al.* performed a study using a wavelet packet decomposition that incorporates the more realistic Poisson noise model of the sinogram data directly into their thresholding estimators [35]. Though using wavelets directly in the reconstruction phase is still an active research area with much promise, we do not pursue this approach in this thesis. Instead, we rely on standard clinical protocols (such as FBP and OSEM) for image reconstruction, then perform denoising through *post-processing* of these images with wavelet-based methods.

4.4 Discrete Dyadic Wavelet Transform

A very important consideration when analyzing PET signals is the selection of an appropriate wavelet expansion set. The recent interest in wavelet research has given a wide variety of possibilities, each of which are suitable for specific areas of application. Though an orthogonal wavelet transform is desirable for some applications (such as compression, due to its compact signal representation), it does not possess good

denoising properties, and is therefore unsuitable for this study. Instead, we seek a translation-invariant wavelet transform appropriate for denoising biomedical signals. Because these signals (which, due to underlying physiological and anatomical factors) are characterized by smooth transitions as opposed to harsh jump discontinuities, it would be unsuitable to choose a basis of sharp-edged functions, such as the step-like Haar representation [13]. Instead, we seek one whose elements better match the features in our PET data. Research using wavelets in medical imaging has shown that *spline-based* wavelet transforms are an appropriate choice, as reviewed in [36].

The criteria above can be satisfied by selecting the *discrete dyadic wavelet transform* (DDWT). The DDWT was introduced by Mallat and Zhong[37] (though they focused on the application of image edge detection), in which they provided transform theory (including a FWT-like filter bank implementation) for one and two dimensional signals. Laine *et al.* performed several denoising and image enhancement studies on biomedical images using the DDWT, among them [38, 39]. In [40], Koren and Laine systematically extended the DDWT theory (and its filter bank implementation) to more than two dimensions, realizing that such a representation can be advantageous in medical imaging applications. Furthermore, the work by Lin *et al.* showed that the DDWT has successfully been used for our class of cardiac PET signals [3].

The DDWT is based on the following sampling of the CWT[10]:

- the scale parameter is sampled along the dyadic sequence $s = 2^j, j \in \mathbb{Z}$, and

- the translation parameter τ is sampled at *all integer locations*, $k \in \mathbb{Z}$. (This differs from the DWT, for which the sampling interval depends on j).

The result is an overcomplete representation with each coefficient subband having the same number of elements as the input signal.

---

[10]See Appendix A for a description of the CWT.

4.4.1 One-dimensional Discrete Dyadic Wavelet Transform (DDWT)

First, we introduce the DDWT for discrete signals of one dimension, $f[n], n \in \mathbb{Z}$, and consider a finite range of scales from 1 to 2^J. Recall that the DWT uses a scaling function to generate a coarse approximation to a signal, accounting for wavelet scales corresponding to $-\infty \leq j \leq 0$. The DDWT uses an analogous construct called a *smoothing function*, $s(x) \in L^2(\mathbb{R})$, for which we define the *smoothing* of an arbitrary function $g(x)$ at scale 2^j (with the operator $\mathcal{S}_{2^j}$) as:

$$\mathcal{S}_{2^j}\{g(x)\} = g(x) * s_{2^j}(x) = g(x) * 2^j s\left(2^j x\right) \tag{4.14}$$

Returning to our discrete input signal, it was proved in [37] that we can write $f[n]$ as the smoothing at scale 1 ($j = 0$) of a non-unique continuous function $f_s(x) \in L^2(\mathbb{R})$, which is subsequently sampled:

$$f[n] = \mathcal{S}_1 f_s(x)\Big|_{x=n,n\in\mathbb{Z}} \tag{4.15}$$

Using this, we define the DDWT of the sequence $f[n]$ (in terms of its "smoothed" function $f_s(x)$) at scale J as the following sets of coefficients indexed by j and k:

$$\left\{ \{\mathcal{S}_{2^J}\{f_s(k+b)\}\}_{k\in\mathbb{Z}}, \{\mathcal{W}_{2^j}\{f_s(k+b)\}\}_{j\in[1,J],k\in\mathbb{Z}} \right\} \tag{4.16}$$

where b is a shift parameter that is determined by the selection of wavelet function and $\mathcal{W}_{2^j}$ is a wavelet operator defined with respect to the mother wavelet function $\psi(x)$:

$$\mathcal{W}_{2^j}\{f(x)\} = f(x) * \psi_{2^j}(x) = f(x) * 2^j \psi(2^j x) \tag{4.17}$$

Like the DWT, the DDWT consists of a group of coefficients representing a low-

pass approximation, along with a set of wavelet coefficient groups at scales 1 to J representing increasingly fine details. To simplify the notation, we introduce the following to represent the DDWT coefficients:

- **Approximation Coefficients:** $\{S_{2^J}[k]\} = \{\mathcal{S}_{2^J} f_s(k+b)\}_{k\in\mathbb{Z}}$

- **Detail Coefficients:** $\{W_{2^j}[k]\} = \{\mathcal{W}_{2^j} f_s(k+b)\}_{j\in[1,J],k\in\mathbb{Z}}$

The mother wavelet that we use in this study is one of a family of wavelets defined in [40] and specified in the frequency domain by:

$$\Psi(\omega) = (j\omega)^r \left(\frac{\sin(\omega/4)}{\omega/4}\right)^{p+r+1} \tag{4.18}$$

A selection with parameters $r = 1, p = 2$ was used with success in another quantitative PET perfusion study [3], so we make a similar choice, giving the following wavelet and smoothing functions:

$$\Psi(\omega) = (j\omega) \left(\frac{\sin(\omega/4)}{\omega/4}\right)^4 \tag{4.19}$$

$$S(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^3 \tag{4.20}$$

Note that this choice of wavelet function can be written as the first derivative of a spline function $\theta(x)$ with Fourier transform:

$$\Theta(\omega) = \left(\frac{\sin(\omega/4)}{\omega/4}\right)^4 \tag{4.21}$$

In practice, the DDWT is calculated with a hierarchical digital filter structure, whose frequency responses must meet a strict set of criteria in the Fourier domain involving $\Psi(\omega)$, $S(\omega)$, and $\Theta(\omega)$ [40]. Details of the filter bank implementation are given in Chapter 5.

4.4.2 Multi-dimensional Discrete Dyadic Wavelet Transform

Having discussed the one-dimensional DDWT, we now consider its extension to multiple dimensions that we require to perform an analysis of dynamic PET data [40].

We will use $\mathcal{S}_{2^j,D}$ to denote a D-dimensional version of the smoothing operator which uses the smoothing function $s(x_1,...,x_D) \in L^2(\mathbb{R}^D)$ to smooth an arbitrary input $g(x_1,...,x_D) \in L^2(\mathbb{R}^D)$:[11]

$$\mathcal{S}_{2^j,D}\{g(x_1,\ldots,x_D)\} = g(x_1,...,x_D) * s_{2^j}(x_1,\ldots,x_D) \qquad (4.22)$$

Note that in this expression, $s_{2^j}(x_1,\ldots,x_D) = 2^j s\left(2^j x_1,\ldots,2^j x_D\right)$ – an extension of the 1D case.

For a D–dimensional DDWT, there are D versions of the wavelet function, indexed by d, $\{^1\psi(x_1,\ldots,x_D),\, ^2\psi(x_1,\ldots,x_D),\, \ldots\, ^D\psi(x_1,\ldots,x_D)\}$, oriented along the different dimensions (*e.g.* for $D = 2$, we have $^1\psi(x_1,x_2)$ - oriented along dimension one - and $^2\psi(x_1,x_2)$ - oriented along dimension two). These define the wavelet operators:

$$^1\mathcal{W}_{2^j,D}\{f(x_1,\ldots,x_D)\} = f(x_1,\ldots,x_D) * {}^1\psi_{2^j}(x_1,\ldots,x_D) \qquad (4.23)$$

$$\vdots$$

$$^D\mathcal{W}_{2^j,D}\{f(x_1,\ldots,x_D)\} = f(x_1,\ldots,x_D) * {}^D\psi_{2^j}(x_1,\ldots,x_D) \qquad (4.24)$$

where $^d\psi_{2^j}(x_1,\ldots,x_D) = 2^j \cdot {}^d\psi\left(2^j x_1,\ldots,2^j x_D\right)$.

We define the DDWT of a D-dimensional discrete signal $f[n_1,\ldots,n_D]$ at scale J

---

[11]The convolution of two functions f_1 and f_2 in d dimensions is given by:
$f_1(x_1,\ldots,x_D) * f_2(x_1,\ldots,x_D) = \int_\infty^\infty \ldots \int_\infty^\infty f_1(\eta_1,\ldots,\eta_D)f_2(x_1 - \eta_1,\ldots,x_D - \eta_D)d\eta_1\ldots d\eta_D$

as the following sets of coefficients:

$$\{\mathcal{S}_{2^J,D}f_s(n_1 + b, \ldots, n_D + b), \{^1\mathcal{W}_{2^j,D}f_s(n_1 + b, \ldots, n_D + b)\},$$

$$\{^2\mathcal{W}_{2^j,D}f_s(n_1 + b, \ldots, n_D + b)\}, \ldots, \{^D\mathcal{W}_{2^j,D}f_s(n_1 + b, \ldots, n_D + b)\}\}$$

where $j \in [1, J]$ and b is a shift parameter that is determined by the selection of wavelet function. Note that this is defined in terms of f's smoothed continuous function, $f_s(x_1, \ldots, x_D) \in L^2(\mathbb{R}^D)$, a multidimensional version of equation 4.14. Again simplifying this notation, we refer to the following DDWT coefficients:

- **Approximation Coefficients:**

$$\{S_{2^J,D}[n_1, \ldots, n_D]\} = \{\mathcal{S}_{2^J,D}f_s(n_1 + b \ldots, n_D + b)\} \qquad (4.25)$$

- **Detail Coefficients:**

$$\{^1W_{2^j,D}[n_1, \ldots, n_D]\} = \{^1\mathcal{W}_{2^j,D}f_s(n_1 + b, \ldots, n_D + b)\} \qquad (4.26)$$

$$\vdots$$

$$\{^DW_{2^j,D}[n_1, \ldots, n_D]\} = \{^D\mathcal{W}_{2^j,D}f_s(n_1 + b, \ldots, n_D + b)\} \qquad (4.27)$$

where $j \in [1, J]$ and $(n_1, \ldots, n_D) \in \mathbb{Z}^D$. Unlike the 1D DDWT, there are D sets of detail coefficients at each scale (one for each dimension), each of which are oriented along that dimension (*e.g.* the $^1W_{2^j,D}$ coefficients will show features oriented along dimension one). To facilitate the implementation of this transform, the multidimensional wavelet functions are chosen as separable products involving the one

dimensional wavelet and smoothing functions given in equations 4.18 and 4.19:

$$^i\psi(x_1,\ldots,x_D) = \psi(x)2^{i-1}\prod_{n=1,n\neq i}^{D} s(2x_n) \tag{4.28}$$

for $i \in [1, D]$. In three dimensions, for example, we have the following wavelet functions:

$$^1\psi(x_1, x_2, x_3) = 4\psi(x_1)s(2x_2)s(2x_3) \tag{4.29}$$

$$^2\psi(x_1, x_2, x_3) = 4\psi(x_2)s(2x_1)s(2x_3) \tag{4.30}$$

$$^3\psi(x_1, x_2, x_3) = 4\psi(x_3)s(2x_1)s(2x_2) \tag{4.31}$$

To illustrate a practical example of the DDWT, Figure 4.6 shows the standard 512x512 *cameraman* image, along with its 3-level, two-dimensional ($D = 2$) DDWT, both approximation and detail coefficients. The following points should be noticed:
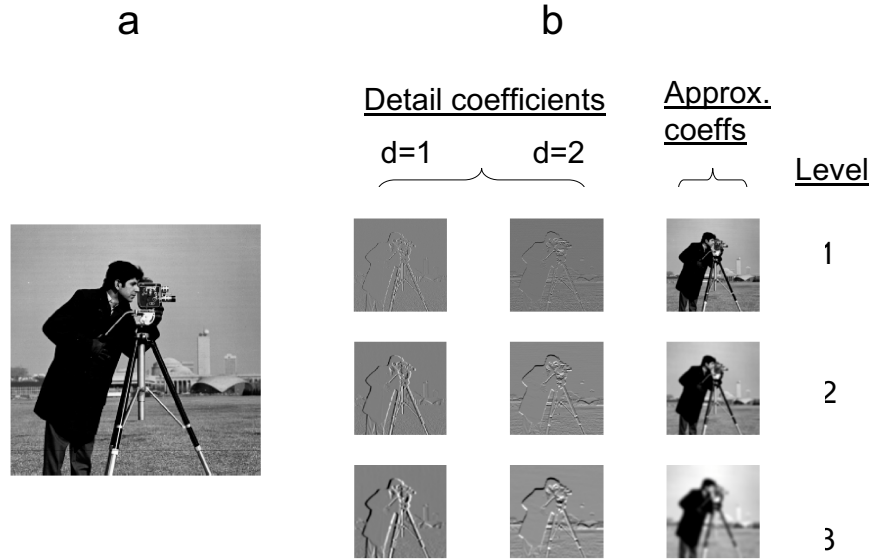
Figure 4.6: a) Cameraman test image and b) the coefficients resulting from a 3-level, 2D DDWT of cameraman test image.

- the features in the detail coefficients associated with dimension one are oriented in the vertical direction,

- the features in the detail coefficients associated with dimension two are oriented in the horizontal direction,

- the size of the features in the detail coefficients increases with the number of decomposition levels,

- the approximation coefficients provide a lower resolution version of the original image. The resolution of this approximation decreases as the number of decomposition levels increases.

With the relevant theory and background in hand, Chapter 5 presents our study investigating the application of wavelet–based denoising methods to ^{82}Rb cardiac PET data.

Chapter 5

Wavelet–based Denoising of

Cardiac PET Data

We now turn our attention to the data of interest - cardiac ^{82}Rb PET images. After first describing the input data (both phantom and clinical) used in this study in Section 5.1, we explain how noise can be characterized in Section 5.2. Section 5.3 provides the details of how the DDWT is implemented as an iterated filter bank. The proposed wavelet-based denoising protocols are fully explained in Section 5.4, and a description of the figures of merit used to measure denoising performance is given in Section 5.5. In Section 5.6 the experimental data are provided and finally, Section 5.7 discusses these results and presents conclusions that can be drawn from this investigation.

In this chapter, an abbreviated notation is used, realizing in all cases that the wavelet basis is the spline-based DDWT described in Chapter 4. Since all denoising methods use three decomposition levels (as explained in Section 5.4), we omit this subscript yielding the following notation that is used in this chapter:

- $\mathcal{W}_D$ – the operator representing a D-dimensional, 3-level DDWT.

- $\mathcal{W}_D^{-1}$ – the operator representing a D-dimensional, 3-level inverse DDWT.

- S_l – the DDWT approximation coefficients at level l.

- ^dW_l – the DDWT detail coefficients at level l, oriented along dimension d.

Notice that we have changed the subscript of Chapter 4 from one indicating *scale*, (*e.g.* 2^j) to one indicating the *level of decomposition* in the filter bank implementation, as this is a more convenient notation to use when explaining the denoising protocols. Furthermore, the dimensionality of the transform (D) appears in the subscripts of the DDWT operators ($\mathcal{W}_D$, $\mathcal{W}_D^{-1}$) only. The approximation and detail coefficients (S_l, ^dW_l) do not include these explicitly – it will be known from the context of the denoising algorithm.

5.1 Description of Input Data

5.1.1 Phantom PET Data

Measuring the effectiveness of a noise reduction algorithm is not straightforward in the nuclear medicine setting. Traditional figures of merit used in denoising studies, such as peak signal to noise ratio (PSNR) and mean squared error (MSE) cannot be calculated for clinical PET data. This is because the noise-free signal is hidden within the body and therefore inaccessible, posing a problem when trying to *objectively* assess denoising performance. This difficulty can be addressed by using a mathematical *phantom*, a computer generated image with known properties that models the biological region of interest. Corrupting the noise-free phantom image with synthesized noise allows us to compare various denoising algorithms because the "ground truth" data are available.

- **2D Image Data** – For 2D image data, we use the phantom image shown in Figure 5.1. This is a 128x128 image with outer ring (myocardium) and inner circle (blood pool) regions of constant intensity on a solid background of

lower intensity. It models a short axis oriented slice of a healthy myocardium. To incorporate the inherent resolution of the PET scanner, the phantom is smoothed with a 2D Gaussian kernel of 10mm FWHM[12].
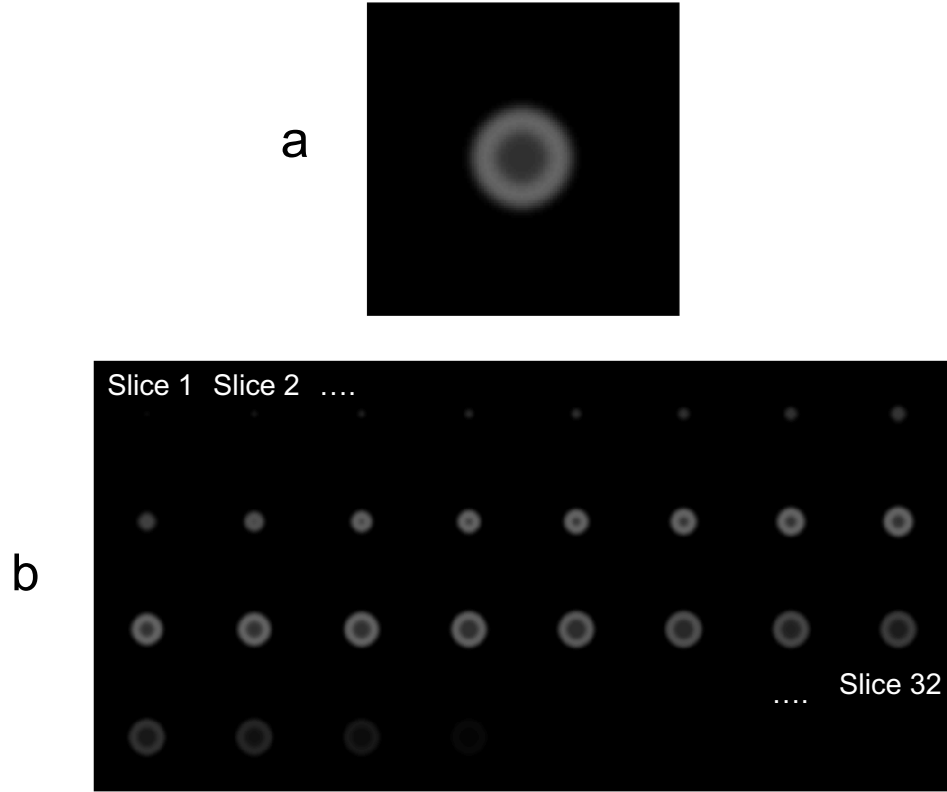
Figure 5.1: Image phantoms used - a) 2D image, b) 3D image slices

- **3D Image Data** – The concept of a phantom image can be easily extended to three dimensions, using 32 simulated slices of image data. Each slice represents a short axis image of a healthy myocardium oriented along the z-axis of the scanner. In order to capture the tapering ellipsoid shape toward the left ventricular apex, the radius of the circles representing the blood pool/inner myocardium boundary and the outer myocardium boundary increases gradually in adjacent slices. The size of each image slice is 128x128 with the myocardium and blood

---

[12]FWHM stands for *full width at half maximum*. It is another measure of the width of the Gaussian, specifying the distance between the points at half the maximum value on either side of the bell curve.

pool regions of constant intensity on a uniform background, and the image slices are equally spaced, forming isotropic voxels. The 3D phantom is shown in Figure 5.1. Like the 2D case, the phantom is smoothed with a Gaussian kernel of 10mm FWHM (this time in 3D) to mimic scanner resolution.

5.1.2 Clinical PET Data

We seek denoising algorithms that are effective for a wide variety of "real world" PET data. To provide a representative sample of the types of clinical images that would benefit from this denoising study, we have selected test cases from OHI with the properties shown in Table 5.1. Note that all of the static data were generated for a single patient (using the conditions specified), and all of the dynamic data were generated for another patient (using the conditions specified).

Data set	Study Conditions	Reconstruction
Static Data		
1	stress	OSEM
2	rest	OSEM
3	stress	FBP
4	rest	FBP
Dynamic Data		
1	stress	OSEM
2	rest	OSEM

Table 5.1: Description of Clinical PET Data

- **Static Image Data** – The static data are in the form of short-axis oriented slices, for healthy subjects with expected uniform myocardium (*i.e.* low risk of CAD). The clinical PET data are generated by the ECAT scanner software using either OSEM or FBP reconstruction as indicated above[6]. OSEM reconstruction uses 16 subsets and 6 iterations, while FBP reconstruction uses a ramp filter. The data are saved in a proprietary binary format called a "ma-

trix" file (stored in 16-bit integers with units of Bq/cc), from which the data
are extracted in a manner appropriate to the denoising algorithm being used.

- **Dynamic Image Data** – The dynamic data are structured in a manner iden-
  tical to that described above, except the slices are transaxial oriented (also for
  healthy subjects), and there are 17 frames of data appearing sequentially in the
  matrix file.

At this point, we must mention that the inherent resolution of the ECAT scanner
used at OHI is 10mm, whereas the pixel sizes of the short axis oriented images in
each dimension are $x = 1.7$mm, $y = 1.7$mm, and $z = 3.7$mm. Therefore, this is a very
low resolution *oversampled* system in which we would not expect to see *meaningful*
image features smaller in size than roughly 5-6 pixels in the x, y directions and 3
pixels in the z direction. Any features present in the raw data that are smaller than
this represent small scale "grains" of noise and are not due to underlying biological
phenomena at this scale.

5.2   Noise Properties of PET Data

The exact characteristics of the noise in our ^{82}Rb PET images are not accurately
known. The positron emission process that is the source of the detected photons is
well characterized, following a Poisson distribution in which a variance estimate is
equal to the number of counts. This distribution does not hold for the PET image
data, however, since image reconstruction and various correction processing (*e.g.*
for attenuation and randoms) must be applied to these data, thereby altering the
noise properties[1]. This assumption is based on the fact that there are a number
of different sources of noise in PET images (as described in Chapter 2), of varying
distributions. The effect of adding these unknown distributions' contribution to a
PET image's pixel activity, will tend toward a Gaussian distribution (as the central

limit theorem states)[5]. Several papers [41, 42] have investigated the noise of both FBP and OSEM image reconstruction methods, which are very different and have the following qualities (as can be seen from Figure 2.7):

FBP reconstruction

- tends to spread noise variance from regions containing high count densities to regions containing low count densities, so the noise variance becomes more uniform throughout the image

- as the FBP filter cutoff frequency decreases, noise become correlated over longer distances

OSEM reconstruction

- in lower count regions of an image, noise variance is much lower than with FBP

- the variance as a function of position closely resembles the noise-free image

- as iteration number increases, noise become correlated over shorter distances

For postprocessing of PET images (either FBP or OSEM), a common assumption is that the noise is Gaussian or uncorrelated [5, 11]. The wavelet–based denoising methods investigated in this thesis assume that the noise between adjacent pixels is uncorrelated (*i.e.* independent of position). An inspection of the noisy PET images (such as those in Figure 2.7) reveals that this is not necessarily the case, as there exist localized "clusters" of noise that seem to be dependent on position. However, we anticipate that the denoising methods will still be effective, based on the study by Lin *et al.* [3] that showed that wavelet-based denoising performance remained satisfactory and robust in the presence of such errors.

One method of estimating the noise PDF's parameters (when only the image data are available) is to select a region in the image that is free of image features and generate the histogram of the pixels in this subimage [13]. For our PET data, we

select a slice in the data set that is far removed from myocardial activity (slice #28 of 32) and compute the histogram of the pixels in a rectangular window. This was done for multiple image files, and representative results are shown in Figure 5.2 (for OSEM) and Figure 5.3 (for FBP).
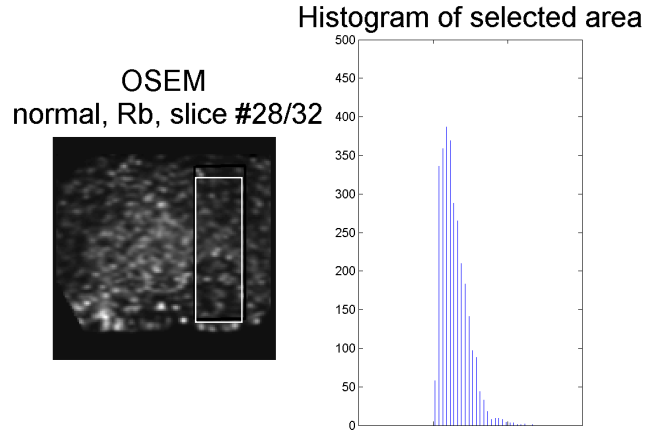

Figure 5.2: Histogram of OSEM noise, using pixels in rectangular window shown
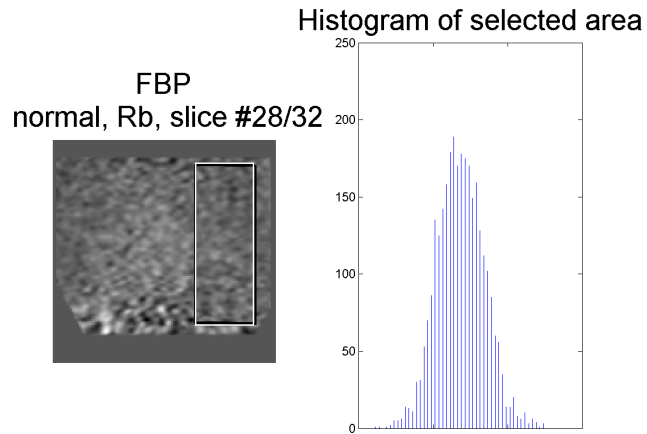

Figure 5.3: Histogram of FBP noise, using pixels in rectangular window shown

Notice that the symmetric histogram of the FBP noise closely resembles a zero mean Gaussian PDF, while that of the OSEM noise is skewed somewhat, more like a

Poisson PDF.

5.2.1 Noise Level Estimation

Wavelet-based denoising methods require the specification of σ_n, the standard deviation of the image noise. Since this is unknown *a priori*, it is necessary to estimate it ($\widehat{\sigma}_n$) from the measured data. This is done using the robust median estimator described in Chapter 4. In our case, a slight modification to the $\widehat{\sigma}_n$ estimation is necessary, since we are not using a wavelet transform for which there is a single subband at the most detailed scale. Instead, the D-dimensional DDWT has detail coefficients spread over D such subbands, so we adopt the following calculation based on the wavelet modulus (see Section 5.3.2)[3]:

$$\widehat{\sigma}_n = \text{Median}(abs(M_1[k]))/0.6745 \qquad (5.1)$$

where k is an index over all coefficient locations at level 1.

At this point, we emphasize one of the key advantages of using wavelet-based denoising over convolution with a fixed width Gaussian kernel. The above discussion illustrates that by using a wavelet representation, we can form an estimate of the PET data noise $\widehat{\sigma}_n$ based on the input data (more specifically in this case, the DDWT coefficients of the input data). This gives a *data-driven* noise estimate that automatically adapts the denoising algorithm to varying noise levels. Because (as we will see in equation 5.9) $\widehat{\sigma}_n$ appears in the expression for the coefficient threshold, the noise estimate ensures that the appropriate amount of smoothing occurs. For example, very noisy input data will give a higher $\widehat{\sigma}_n$ estimate, thereby increasing the threshold and providing more aggressive denoising in images that require it. Such behaviour is not possible with the Gaussian approach, which uses a fixed width kernel regardless of the noise level of the input data.

5.2.2 Noise Simulation

We have already remarked that when performing phantom studies, the phantom data need to be corrupted with noise. For a noise type like additive white Gaussian noise (AWGN), this is a trivial task since software that reliably produces samples from the desired distribution is available. The noise properties of our PET images are not precisely known, so generating noise samples in a similar manner from a complicated unknown distribution is not an option. Of course, in order that our denoising algorithms be effective for clinical PET data, we seek a simulated noise realization that is as close as possible to that generated from an actual scanner. To accomplish this, we choose a slice of the data set that is free of myocardial activity, situated at the edge of the z axis field of view (shown in Figure 5.4). In the absence of
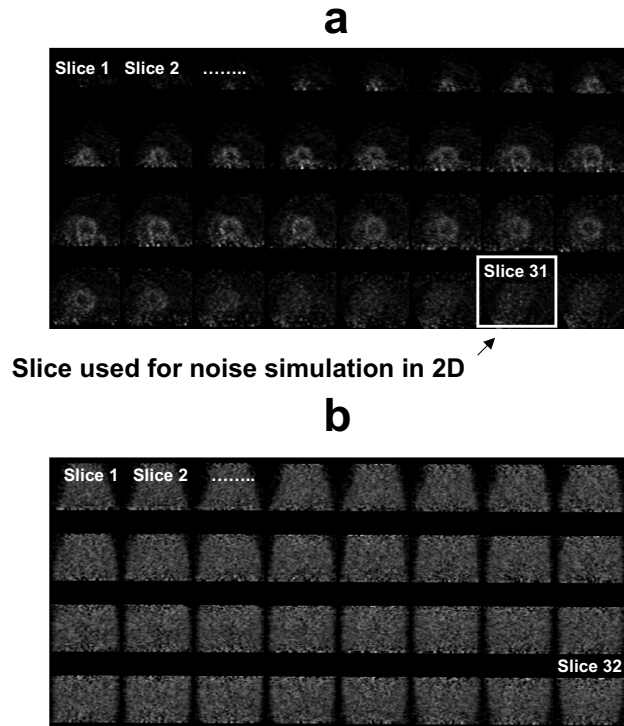

Figure 5.4: a) Slices of clinical PET data set (OSEM reconstruction), with the slice used for simulation of 2D noise highlighted. b) 3D noise-only slices of water cylinder scan (OSEM reconstruction). The same is done for FBP data.

signal, this image is assumed to be due to noise only, and is fused with the phantom data to create the simulated noisy image needed[43].

The above approach works in two dimensions, but the situation is different in three dimensions, where we require 32 slices of "noise only" data. A naive approach might be to simply duplicate the noisy image for all 32 slices, but this would introduce a false correlation between slices, potentially biasing the results. A method of generating 32 slices of simulated PET noise that realistically models the correlation between slices is with the use of a water cylinder phantom. A 20cm diameter, 20cm long water cylinder was injected with an amount of ^{82}Rb radioactivity similar to that used for a static clinical image, providing a uniform test signal free of any localized uptake. Thus, when the images from these data are reconstructed using the same parameters and orientation as the static image, we arrive at a valid estimate of 32 "noise only" slices. As in 2D, this image is assumed to be due to noise only, and is fused with the phantom data to create the simulated noisy image.

5.3 Implementation Details for Discrete Dyadic Wavelet Transform

5.3.1 One–dimensional DDWT

As mentioned in the previous chapter, an efficient DDWT implementation is possible using a filter bank structure. A three level decomposition is illustrated in Figure 5.5, showing how the approximation and detail coefficients at various levels are obtained through iterated filtering of the input signal. The filter specifications at each level are more complicated than that of the DWT, being defined in terms of the following three filter transfer functions. Derivation of this filter bank structure is beyond the scope of this thesis, but the reader is referred to [40] for details.

a

Input
f [n]

$G_s(\omega)$ $^1W_1$

$H_s(\omega)$ $G_s(2\omega)$ $^1W_2$

$H_s(2\omega)$ $G_s(4\omega)$ $^1W_3$

$H_s(4\omega)$ S_3

b

$^1W_1$ $K_{-s}(\omega)$

$^1W_2$ $K_{-s}(2\omega)$ Output
f[n]

$^1W_3$ $K_{-s}(4\omega)$ $H_s^*(\omega)$

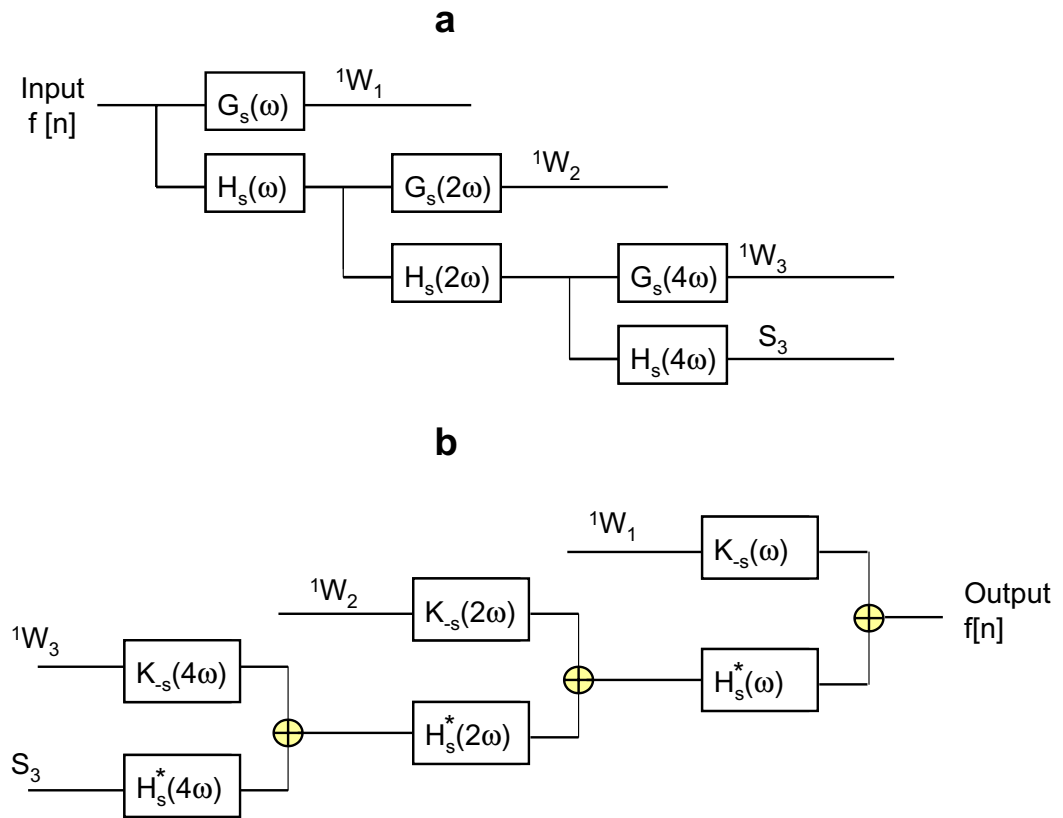S_3 $H_s^*(4\omega)$ $H_s^*(2\omega)$

Figure 5.5: 1D DDWT filter bank structure, 3 levels - a) Analysis, b) Synthesis

$$H(\omega) = e^{j\omega/2}(\cos(\omega/2))^3 \tag{5.2}$$

$$G(\omega) = e^{j\omega/2}4j\sin(\omega/2) \tag{5.3}$$

$$K(\omega) = -\frac{1}{16}\left(e^{-j\omega}G(\omega)\right)\left(1 + (\cos(\omega/2))^2 + (\cos(\omega/2))^4\right) \tag{5.4}$$

Note that $H(\omega)$ describes a low pass filter, while $G(\omega)$ and $K(\omega)$ are high pass filters. All are finite impulse response (FIR) filters with impulse response coefficients appearing in Table 5.2. In the filter bank structure, the s subscript represents a shift in the spatial domain – in general, $F_s(\omega) = e^{-j\omega s}F(\omega)$. At levels two and above, the filters appearing in that stage are dilations of the level one filters by a factor 2^{l-1}. (*e.g.* at level 2, $H(\omega)$ becomes $H(2^{2-1}\omega) = H(2\omega)$, and at level 3, $G(\omega)$ becomes $G(2^{3-1}\omega) = G(4\omega)$). In the spatial domain, this can be implemented by upsampling the filter response by a factor of 2^{l-1} (*i.e.* inserting $2^{l-1}-1$ zeros between samples of the sequences given in Table 5.2). This method is known in the literature as the *à trous* algorithm, and was introduced by Mallat [44]. Notice that the complex conjugate of the analysis $H(\omega)$ filter is used in the synthesis portion, and that the structure in Figure 5.5 does not have downsampling in the analysis portion (nor upsampling in the synthesis branches), as is the case with the DWT. This gives the translation-invariant transform required for denoising purposes.

5.3.2 D–dimensional DDWT

The multidimensional DDWT is also implemented with an efficient algorithm[40]. The general filter bank structure for a D-dimensional DDWT (this time using two levels) is shown in Figure 5.6 (analysis) and Figure 5.7 (synthesis). Examining these, we notice that unlike the 1D case, the filter bank branches involve cascades of multiple filters.

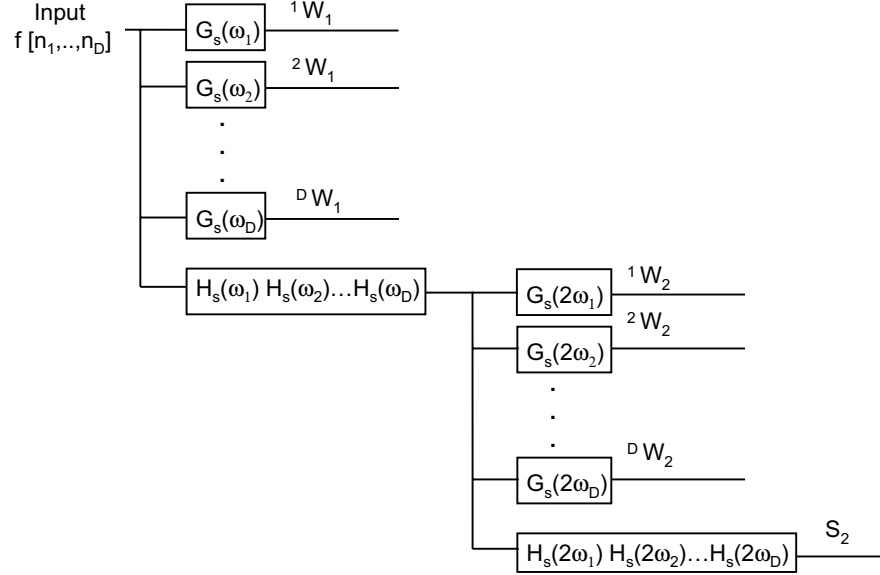In the analysis portion, the low pass filtering is now performed using a cascade

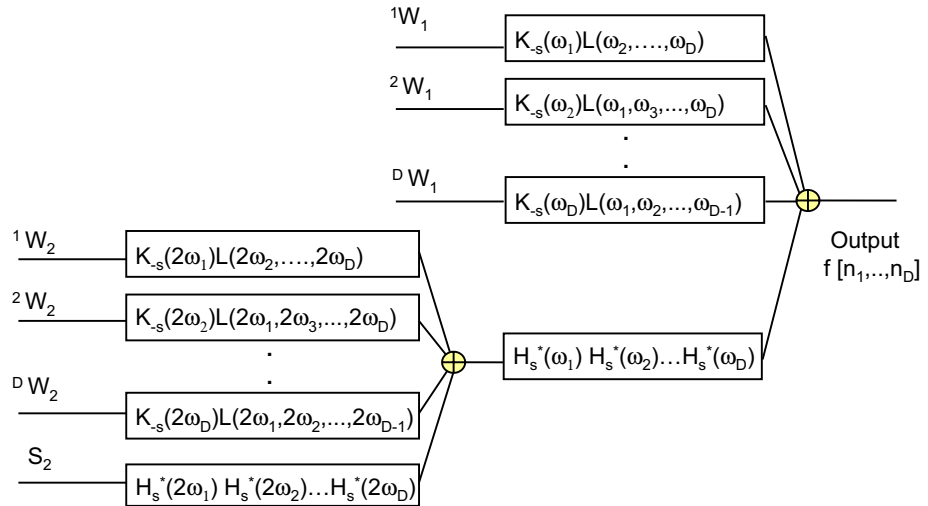**Figure 5.6:** D-dimensional DDWT analysis filter bank structure

Figure 5.7: D-dimensional DDWT synthesis filter bank structure

n	$\mathbf{h[n]}$	$\mathbf{g[n]}$	$\mathbf{k[n]}$	$\mathbf{l_2[n]}$	$\mathbf{l_3[n]}$
-4	0	0	0	0	0
-3	0	0	0	0.0078125	0.015625
-2	0.125	0	-0.0078125	0.046875	0.09375
-1	0.375	2	-0.0546875	0.1171875	0.234375
0	0.375	-2	-0.171875	0.65625	0.8125
1	0.125	0	0.171875	0.1171875	0.234375
2	0	0	0.0546875	0.046875	0.09375
3	0	0	0.0078125	0.0078125	0.015625
4	0	0	0	0	0

Table 5.2: Filter coefficients for implementation of spline-based DDWT

of separable filters (we use the $\omega_1, \omega_2, \ldots \omega_D$ frequency variables to represent filtering along the different dimensions of the input signal). For example, with $D = 3$, this separable configuration requires filtering the data sequentially along the first, second, and third dimensions with the same filter kernel $H(\omega)$. In addition, there are now D high pass filter branches, each consisting of a single filter $G(\omega)$ that filters the data along each of the dimensions separately. Filters at level $l \geq 2$ are modified in the same way as in the 1D case - with frequency dilation by a factor of 2^{l-1} (which can be "*à trous*" implemented in the spatial domain).

The synthesis portion is slightly more complicated. The branches that perform the filtering of detail coefficients consist of a single dimensional $K(\omega)$ filter in cascade with a multidimensional L_D filter:

For $D = 2$,

$$L_2(\omega) = \frac{1}{2}\left(1 + |H(\omega)|^2)\right)$$

For $D = 3$,

$$L_3(\omega_1, \omega_2) = \frac{1}{3} + \frac{1}{6}\left(|H(\omega_1)|^2 + |H(\omega_2)|^2\right) + \frac{1}{3}|H(\omega_1)H(\omega_2)|^2$$

The L_3 filter is not separable, but can be reformulated as a combination of separable

functions:

$$L_3(\omega_1, \omega_2) = \frac{1}{3} |H(\omega_1)|^2 |H(\omega_2)|^2 + \frac{1}{6} |H(\omega_1)|^2 + \frac{1}{6} |H(\omega_2)|^2 + \frac{1}{3}$$

The synthesis branches that filter the approximation coefficients are filtered in all dimensions by $H^*(\omega)$. All filter coefficients appear in Table 5.2.

Wavelet Modulus – For the multidimensional DDWT, it is also necessary to define the *wavelet modulus*, M_l, which is computed at each position in decomposition level l as follows:

$$M_l[n_1, \ldots, n_D] = \sqrt{(^1W_l[n_1, \ldots, n_D])^2 + \ldots + (^DW_l[n_1, \ldots, n_D])^2} \qquad (5.5)$$

This quantity can be interpreted as the length of a vector formed from the D detail coefficients at position $[n_1, \ldots, n_D]$ in level l,

$$^l\mathbf{W}[n_1, \ldots, n_D] = [^1W_l[n_1, \ldots, n_D] \ldots ^D W_l[n_1, \ldots, n_D]]^T \qquad (5.6)$$

with direction defined by the unit vector:

$$^l\mathbf{w}[n_1, \ldots, n_D] = \frac{1}{M_l[n_1, \ldots, n_D]} [^1W_l[n_1, \ldots, n_D] \ldots ^D W_l[n_1, \ldots, n_D]]^T \qquad (5.7)$$

As we see in the next section, the wavelet modulus is used to determine denoising threshold values.

5.4 Description of Denoising Protocols

5.4.1 Development Environment

The simulations have been performed in the Matlab 6.5 numerical software environment, using functions from the Image Processing Toolbox version 3.2. This environment allows rapid algorithm development, verification, and visualization tools [45, 46]. A graphical user interface (GUI) was implemented in order to facilitate the quick evaluation of parameter changes on the denoising results. A screenshot of this GUI is shown in Figure 5.8.

5.4.2 General Denoising Procedure and Notation Used

As introduced in Chapter 4, we review the overall process of wavelet-based denoising below. We consider both two and three dimensional input data, so each method will use either the two or three dimensional DDWT ($D = 2, 3$).

1. **Forward Wavelet Transform**

   The forward transform (in our case, the D-dimensional, 3-level DDWT) of the noisy input data f_{noisy} is calculated, $\mathcal{W}_D\{f_{noisy}\}$, generating a set of approximation coefficients at level 3, and D sets of detail coefficients at each of levels $l = 1, 2, 3$:

   - **Approximation coefficients at level** 3: S_3

   - **Detail coefficients at level** $l \in 1, 2, 3$: $\left\{^1W_l, {}^2W_l, \ldots, {}^D W_l\right\}$

   Three levels of wavelet decomposition were deemed to be satisfactory for this study. Empirically, we found that using greater than three levels increased computation time but did not improve results.
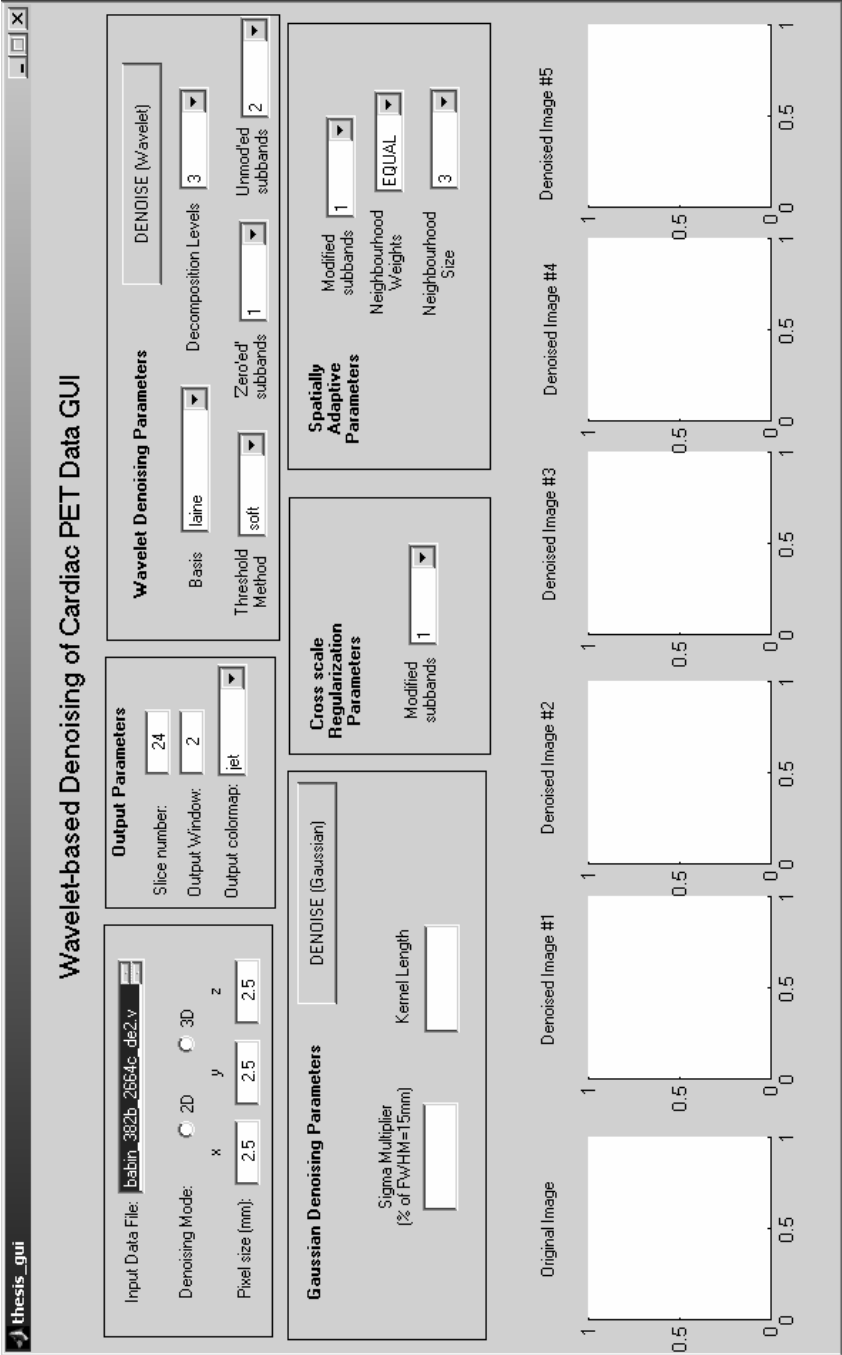
2. **Coefficient Modification**

Figure 5.8: Screenshot of "Wavelet-based Denoising of Cardiac PET Data" GUI

The DDWT detail coefficients are modified according to a thresholding scheme that suppresses the noise contribution in the wavelet domain. The approximation coefficients remain unchanged. We denote the modified, or *denoised*, coefficients with a tilde (˜):

- **Denoised approximation coefficients at level** 3: S_3

- **Denoised detail coefficients at level** $l \in 1, 2, 3$: $\left\{ {}^1\tilde{W}_l, {}^2\tilde{W}_l, \ldots, {}^D\tilde{W}_l \right\}$

3. **Inverse Wavelet Transform**

The inverse transform of the denoised coefficients is calculated to reconstruct the denoised image, $f_{denoised} = \mathcal{W}_D^{-1}\{\tilde{W}\}$ (where $\tilde{W}$ refers to all subbands of denoised coefficients at all levels above).

5.4.3 Details of Denoising Protocols

In this study, we investigate different ways of modifying the detail coefficients at various decomposition levels, and assess the effect that these methods have on the denoised output. The techniques adopted are among the most recent in the wavelet denoising literature, reviewed in Chapter 4:

- spatially adaptive (SA) thresholding

- cross scale regularization (CSR)

For a given overall denoising method (which we call a *denoising protocol*), we modify the detail coefficients using a combination of the above methods, as detailed in Table 5.3.

Please note:

1. When the detail coefficients at a given decomposition level are modified using a certain method, all D subbands at that level are modified in the same manner.

Denoising Protocol	Detail coefficient levels set to zero	Detail coefficient levels unmodified	Detail coefficient levels modified by SA thresholding	Detail coefficient levels modified by CSR
Denoising protocols using inclusion or deletion of entire subbands				
1	1	2 3	–	–
2	1 2	3	–	–
3	1 2 3	–	–	–
Denoising protocols using SA				
4	–	2 3	1	–
5	1	3	2	–
6	–	3	1 2	–
7	1	–	2 3	–
8	1 2	–	3	–
9	–	–	1 2 3	–
Denoising protocols using CSR				
10	–	2 3	–	1
11	–	1 3	–	2
12	1	3	–	2
13	–	3	–	1 2
Denoising protocols using both SA and CSR				
14	–	3	2	1
15	–	–	2 3	1
16	1	–	3	2
17	–	–	3	1 2
Gaussian denoising - STATIC studies				
18	*FWHM = 5 mm*			
19	*FWHM = 10 mm*			
20	*FWHM = 15 mm*			
21	*FWHM = 20 mm*			
Gaussian denoising - DYNAMIC studies				
18	*FWHM = 12 mm*			

Table 5.3: Denoising protocols used in this thesis

2. In some of these protocols, an entire level of coefficients is set to zero. This decision is based on the observation that at the finest scales, noise predominates and it may be better to discard the entire level's effect in the reconstruction.

3. In some protocols, we leave entire levels of coefficients unmodified.

4. In all cases, the approximation coefficients contain important low frequency content and are therefore not modified.

5. The denoising protocols are performed using both 2D and 3D DDWT as follows:

   Static Data

   - In the 2D case, each short-axis oriented slice undergoes 2D denoising. This process is repeated independently for all slices in the volume.

   - In the 3D case, the entire volume of data undergoes 3D denoising at once.

   Dynamic Data

   - In the 2D case, each transaxial oriented slice undergoes 2D denoising. This process is repeated independently for all slices in the volume, and subsequently on all frames in the study.

   - In the 3D case, the entire volume of data undergoes 3D denoising at once. This process is repeated independently for all frames in the study.

6. Protocols 18-21 have been included in order to compare the results our wavelet–based techniques with those from GLPF smoothing of varying width. The smoothing filter is a 2D Gaussian for 2D denoising (slice-by-slice) and a 3D Gaussian for 3D denoising (entire volume at once). Note that protocol 18 is defined differently for static and dynamic studies.

Having presented the denoising protocols with this approach, we need not provide repetitive descriptions of algorithm details in each case. It suffices to describe how

a given modification method is performed on a single arbitrary DDWT coefficient subband (at level l in dimension d) to yield the denoised coefficients. *This is repeated for each subband and for each level in the DDWT representation using the information in Table 5.3.* After all subbands have been processed, reconstruction of the denoised image from the modified coefficients is done using the appropriate inverse DDWT (either 2D or 3D). The description of the different modification methods follows.

Spatially Adaptive Thresholding

The following steps are performed in order to denoise the detail coefficients ^dW_l in the subband being processed (level l, dimension d) by spatially adaptive thresholding [47]:

1. Estimate the noise level $\widehat{\sigma}_n$ from the modulus of the noisy level 1 coefficients using equation 5.1.

2. Calculate the adaptive threshold value at each coefficient location in the subband. To do this, we model each coefficient as a random variable with a zero mean generalized Gaussian distribution, and estimate the value of the distribution's variance[24]. At a given coefficient location in the subband, we call this parameter $\sigma[n_1, \ldots, n_D]$. We compute the variance using an equally weighted average of its neighbouring coefficients[48]:

$$\sigma^2[n_1, \ldots, n_D] = \frac{1}{N} \left(\sum_{[m_1, \ldots, m_D] \in \mathcal{N}} {}^dW_l[n_1, \ldots, n_D]^2 \right) - \widehat{\sigma}_n^2 \qquad (5.8)$$

where $\mathcal{N}$ is a neighbourhood of N pixels around the coefficient currently being processed. We then use this value in the following expression for the spatially adaptive threshold:

$$\eta[n_1, \ldots, n_D] = \frac{\widehat{\sigma}_n^2}{\sigma[n_1, \ldots, n_D]} \qquad (5.9)$$

For this study, $\mathcal{N}$ was selected as a 3x3 (for 2D) or 3x3x3 (for 3D) neighbourhood, a size that was determined empirically.

3. Finally, perform soft thresholding of the noisy coefficients with the spatially adaptive threshold value to obtain the denoised coefficients:

$$^{d}\tilde{W}_l[n_1, \ldots, n_D] = \mathcal{T}^{soft}_{\eta[n_1, \ldots, n_D]} \left\{ ^{d}W_l[n_1, \ldots, n_D] \right\} \tag{5.10}$$

Cross Scale Regularization (CSR)

The following steps are performed in order to modify the detail coefficients $^{d}W_l$ in the subband being processed (level l, dimension d) by cross scale regularization. In this algorithm, the coefficient modification is performed using the wavelet modulus instead of separately on the individual wavelet vector components[26]. This is done in order to preserve the direction of the wavelet vector $\mathbf{w}$ (the gradient direction), in which the signal changes the most rapidly.

1. At each coefficient position, compute the wavelet modulus of level l:

$$M_l[n_1, \ldots, n_D] = \sqrt{(^{1}W_1[n_1, \ldots, n_D])^2 + \ldots + (^{D}W_1[n_1, \ldots, n_D])^2} \tag{5.11}$$

and the unit vector defining its direction:

$$\begin{aligned} \mathbf{w}_l[n_1, \ldots, n_D] &\triangleq \frac{1}{M_l[n_1, \ldots, n_D]} [^{1}W_l[n_1, \ldots, n_D] \ldots ^{D}W_l[n_1, \ldots, n_D]]^T \\ &= [^{1}w_l[n_1, \ldots, n_D] \ldots ^{D}w_l[n_1, \ldots, n_D]]^T \end{aligned}$$

2. Calculate the wavelet modulus at the next coarsest level, $(l+1)$, for each coefficient position and normalize this to within the range 0–1 by dividing by the

maximum value of the wavelet modulus at level $(l + 1)$:

$$M_{l+1}^{norm}[n_1, \ldots, n_D] = \frac{M_{l+1}[n_1, \ldots, n_D]}{max\{M_{l+1}\}} \quad (5.12)$$

Note that if the coefficients at level $(l + 1)$ are being modified as part of the denoising protocol, *that modification should be completed before performing this step.*

3. Find the modified wavelet modulus of the level l coefficients, ${}^l\tilde{M}[n_1, \ldots, n_D]$ by the following pointwise multiplication:

$$\tilde{M}_l[n_1, \ldots, n_D] = M_{l+1}^{norm}[n_1, \ldots, n_D] \cdot M_l[n_1, \ldots, n_D] \quad (5.13)$$

4. Finally, generate the modified detail coefficients at each position by projecting the vector defined by $\tilde{M}_l$ and $\mathbf{w}_l$ onto dimension d:

$$^d\tilde{W}_l[n_1, \ldots, n_D] = M_l[n_1, \ldots, n_D] \cdot \mathbf{w}_l[n_1, \ldots, n_D] \quad (5.14)$$

5.4.4 Signal Extension at Boundaries

Since our DDWT filter bank implementation uses spatial convolution with fixed length filter kernels, we must extend the input signal at the boundaries in each dimension in an appropriate manner. Common approaches in image denoising include zero extension, periodic extension, and mirror extension [13]. Because the first two methods, in general, have the possibility of introducing abrupt changes at the edges, we do not consider these. Instead, we choose mirror extension [49, 40].

5.5 Figures of Merit

From the outset, we emphasize that the denoising protocols ultimately deemed to be of the highest quality will be the ones that produce denoised output that is the most *diagnostically useful in a clinical setting*. The use of phantom images (and the numerical figure of merit that they provide) is a tool to help us toward that end, but not an end in itself.

5.5.1 Phantom Data

1. **Mean–Squared Error (MSE)**

   In this thesis, we use the MSE between the noisy and denoised images as the numerical figure of merit. It is defined as:

   $$MSE = \frac{1}{N} \sum_{i=1}^{N} (f_{denoised}[i] - f_{phantom}[i])^2 \tag{5.15}$$

   where N is the total number of values in the denoised data, $f_{denoised}$ is the denoised image and $f_{phantom}$ is the (known) phantom image. Clearly, a better denoised image will tend to give a lower MSE value.

2. **Visual Assessment**

   Because minimizing the MSE does not necessarily produce the *best* result, visual assessments of these denoised images are necessary. Important criteria include clear delineation of the myocardium region at the inner (blood pool) and outer boundaries, uniformity of activity in the myocardium and in the blood pool, and suppression of undesired activity in the background region. Note that for phantom images, this subjective assessment can easily be done since we know the true image, alongside which we can directly compare the denoised result.

5.5.2 Clinical Data

For clinical data, the denoising results must be evaluated differently. Depending on whether the study in question is a static or dynamic one, these measures are explained below:

1. **Visual Assessment – (used for static studies)**

   Criteria similar to those described above are used when evaluating clinical images. However, this approach requires the expertise of a trained PET technician or cardiologist, who are able to distinguish a diagnostically significant feature of interest (*e.g.* a perfusion defect, papillary muscle attached to the aortic valve at the LV outlet) from a possible artifact of the denoising process.

2. **Coefficient of Determination** (R^2) **– (used for dynamic studies)**

   For dynamic clinical PET data, we can evaluate the effectiveness of the denoising protocol with quantities calculated by OHI's tracer kinetic model software. When the modeling has completed (thereby determining flow parameters), a quantity called the *coefficient of determination*, or R^2, measures how well the compartmental model fits the measured PET data at the solution parameters. R^2 is a calculation based on the sum of the distances from the measurement points to the best fit curve of the model, and can vary between 0 and 1 (a value of 1 corresponds to perfect agreement between model and measurements), so an effective denoising technique should tend to give a higher R^2 value.

3. **Normalized Standard Deviation of** K_1 $(SD_{K1,norm})$ **– (used for dynamic studies)**

   This quantity measures the "stability" of the solution, and it is calculated using the diagonal entries of the covariance matrix output by the model. A small SD_{K1} represents a very steep convergence, giving a solution that is insensitive

to small differences in parameter changes, clearly a desirable property. In order to reduce bias that might be introduced by the K_1 parameter, we use as our figure of merit a normalized value, $SD_{K1,norm} = SD_{K1}/K_1$.

Since model fitting is performed at *each sector in the cardiac polar map*, we take the *median* of the above quantities, in order to minimize the effect of data outliers.

5.6 Experimental Results

In this section, we present the results of applying the denoising protocols described above on ^{82}Rb cardiac PET data.

5.6.1 Phantom Data

MSE Results

- **2D** – In Figure 5.9, we plot the MSE values resulting from applying the 2D denoising protocols when the 2D phantom is corrupted with the indicated noise type.

- **3D** – In Figure 5.10, we plot the MSE values resulting from applying the 3D denoising protocols when the 3D phantom is corrupted with the indicated noise type, using slice 20/32 in the MSE calculation.
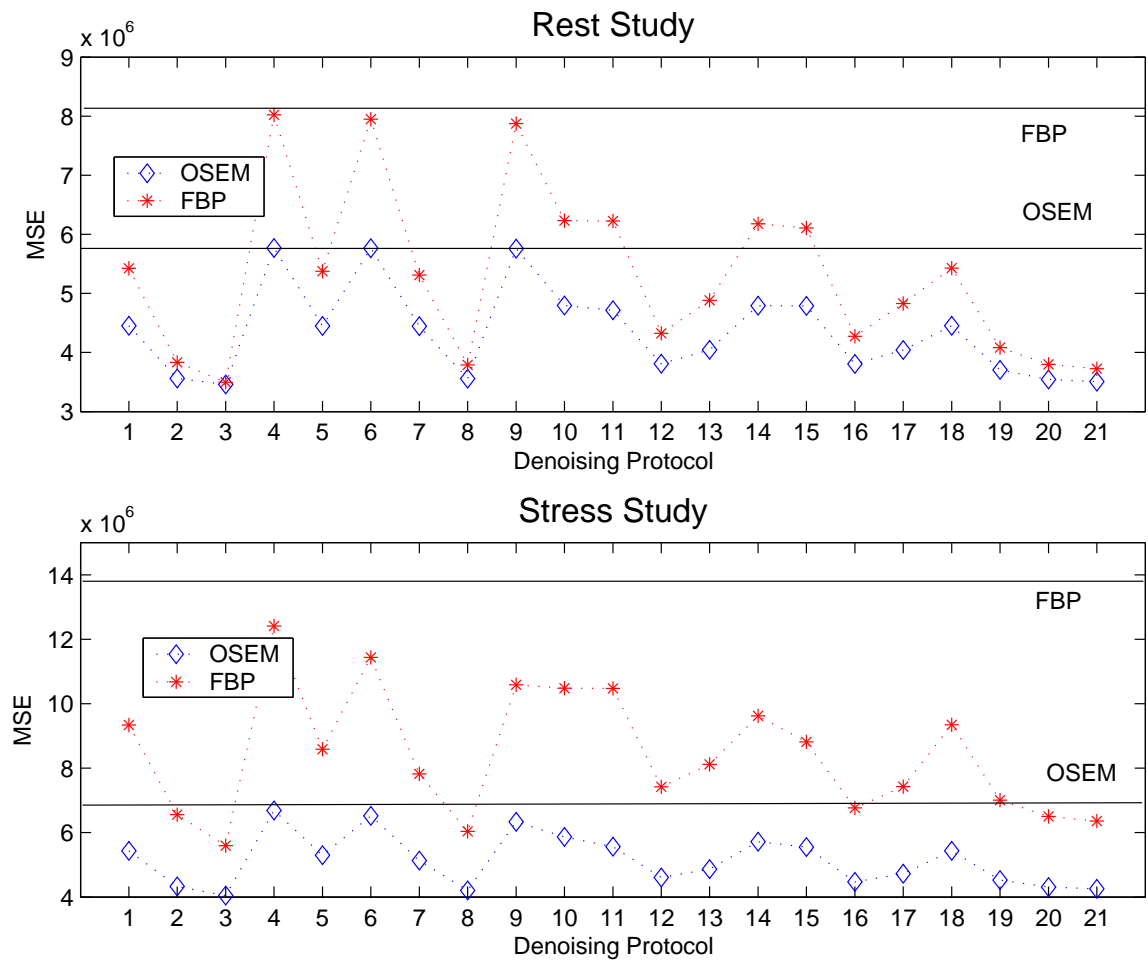
Figure 5.9: MSE results of denoising protocols on 2D phantom. The MSE of the noisy images are indicated with the horizontal lines.
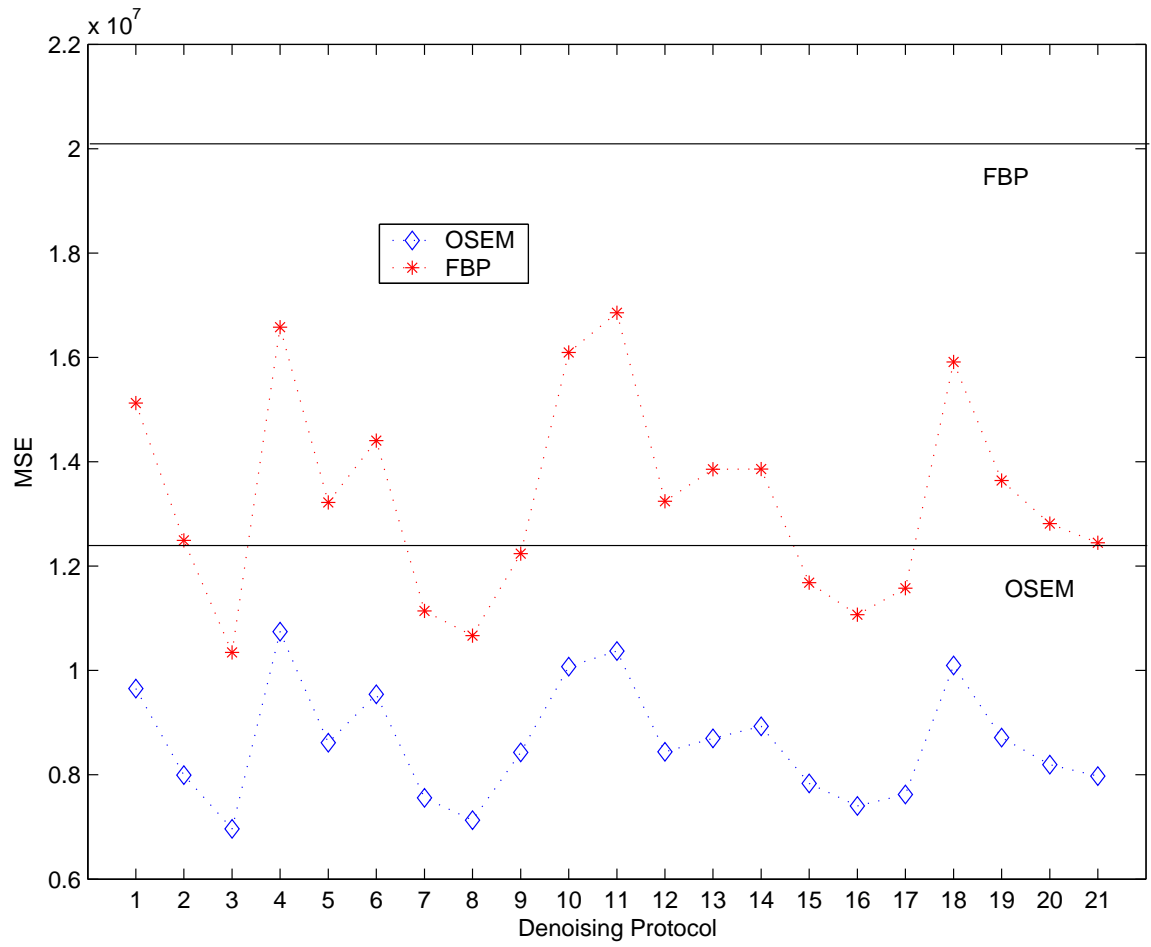
Figure 5.10: MSE results of denoising protocols on 3D phantom. The MSE of the noisy images are indicated with the horizontal lines.

Results from Visual Assessment

Some selected visual results of processing the corrupted phantom images with the denoising protocols are presented as follows:

- Figure 5.11 — 2D phantom image corrupted with OSEM stress noise

- Figure 5.12 — 2D phantom image corrupted with FBP stress noise

- Figure 5.13 — 3D phantom image corrupted with OSEM noise

- Figure 5.14 — 3D phantom image corrupted with FBP noise

All figures that show denoised images in this thesis appear in colour, using a "jet" colour scale in which blue represents the minimum value in the image and red represents the maximum value in the image. In the 3D phantom case, we have chosen to display the mid–ventricular slice # 20/32, due to its similarity to the 2D phantom image. Notice that in each figure, the original noisy image appears in the lower right corner.
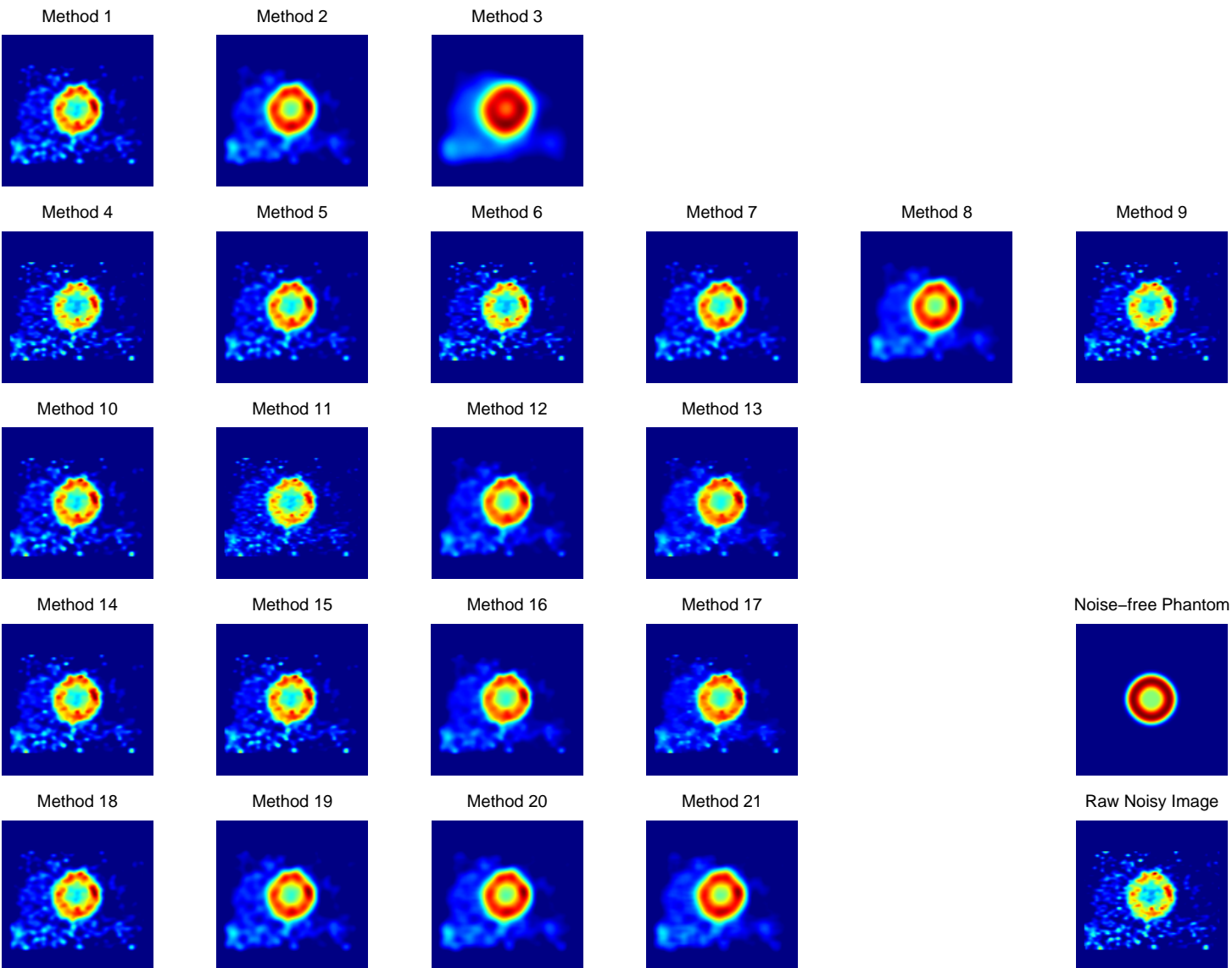
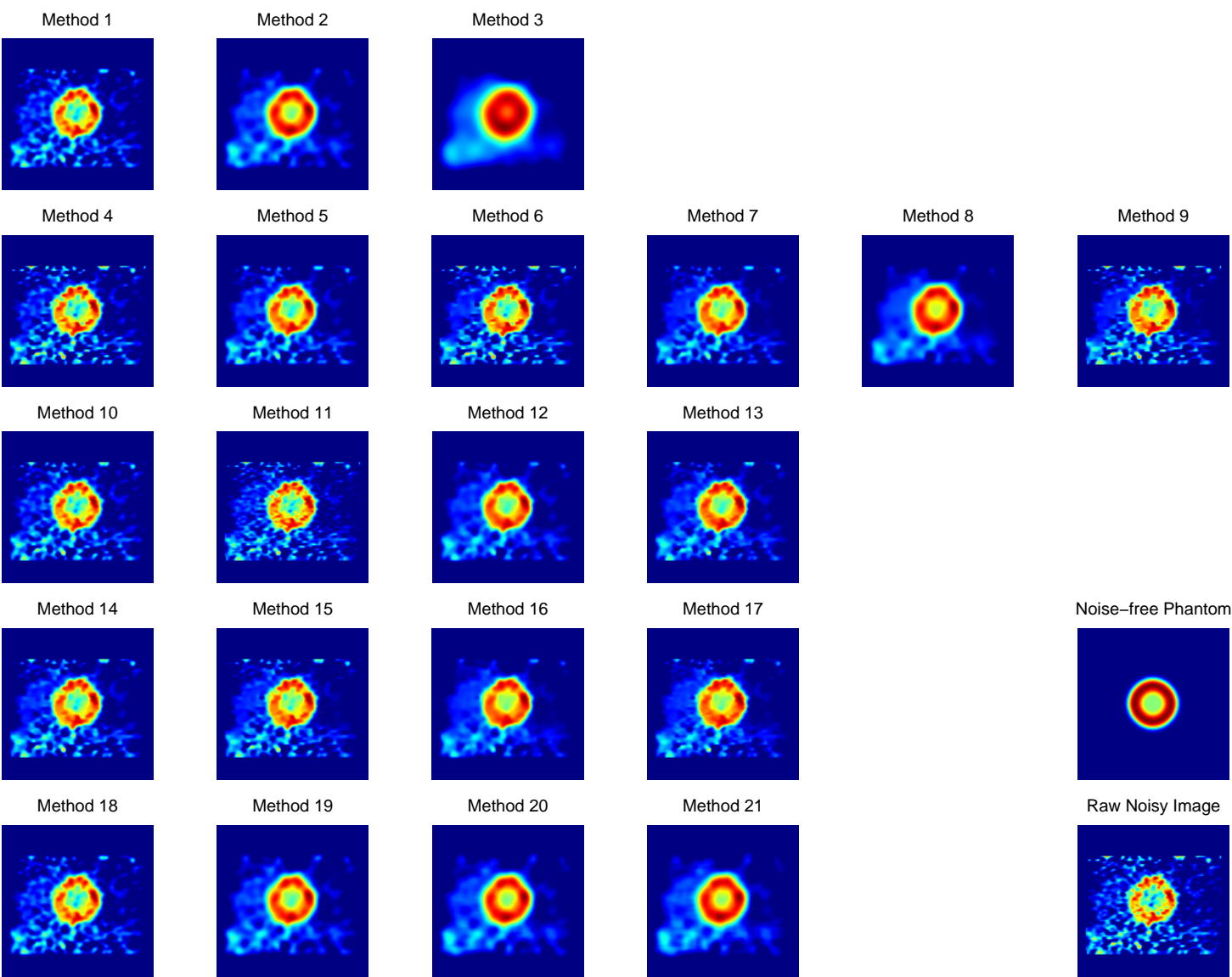**Figure 5.11:** Visual results of denoising protocols on 2D phantom, simulated OSEM stress noise

Figure 5.12: Visual results of denoising protocols on 2D phantom, simulated FBP stress noise
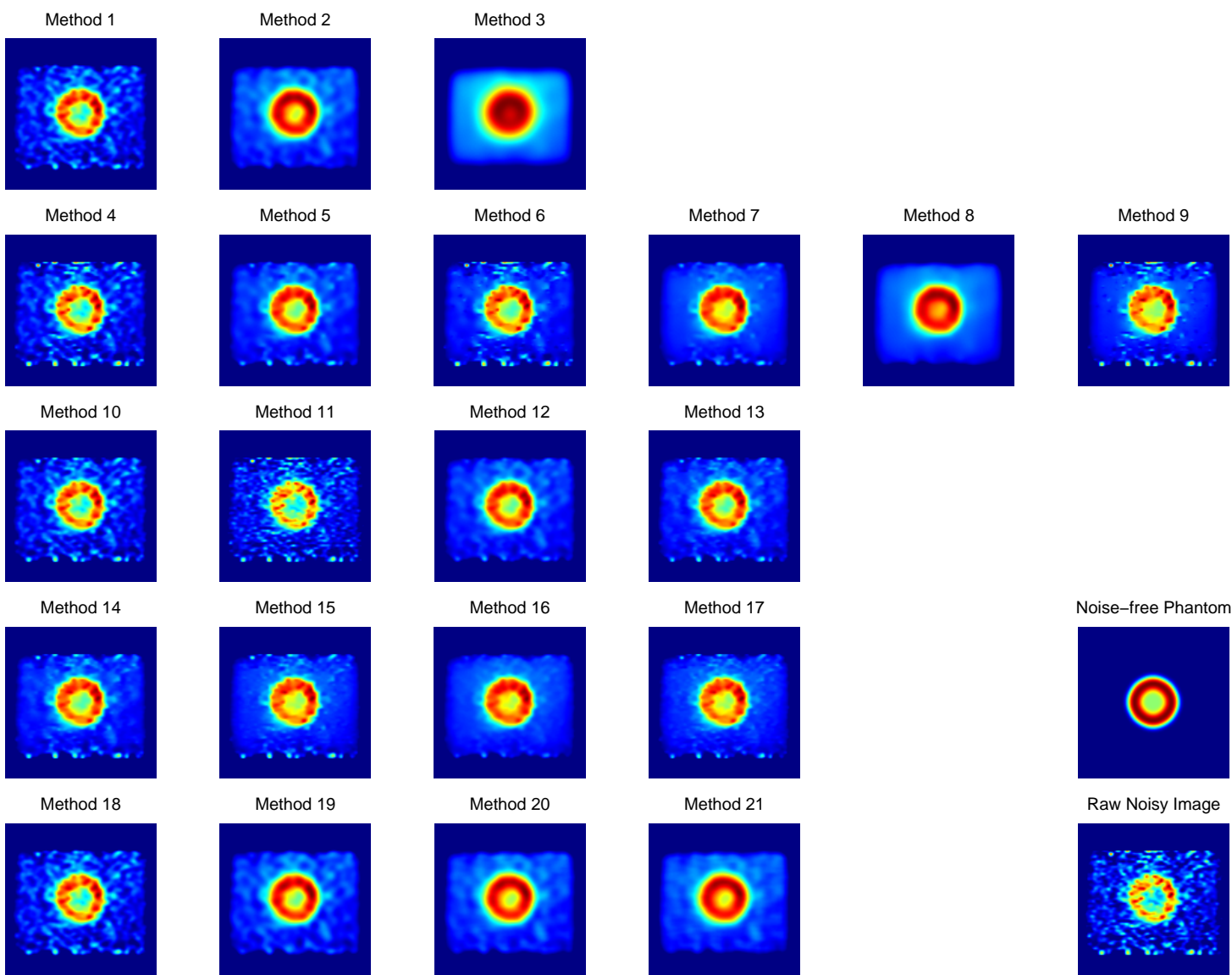
Figure 5.13: Visual results of denoising protocols on 3D phantom, simulated OSEM noise, slice 20 shown
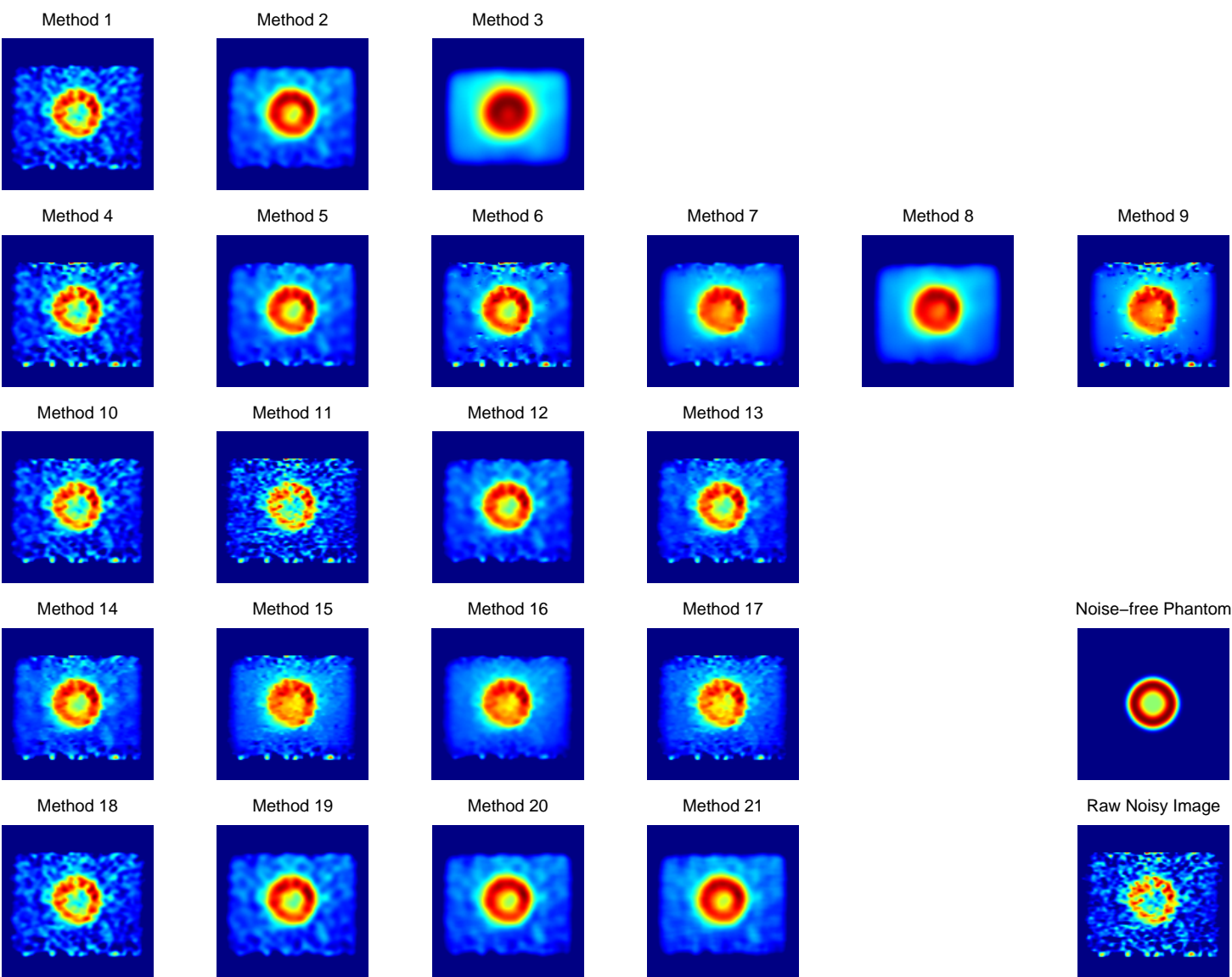
Figure 5.14: Visual results of denoising protocols on 3D phantom, simulated FBP noise, slice 20 shown

5.6.2 Clinical Data

Static Data – Results of Visual Assessment

The following figures show selected visual results of processing the static clinical PET images with the denoising protocols:

- Figure 5.15 — 2D denoising methods, OSEM stress image

- Figure 5.16 — 2D denoising methods, FBP stress image

- Figure 5.17 — 3D denoising methods, OSEM stress image

- Figure 5.18 — 3D denoising methods, OSEM rest image

- Figure 5.19 — 3D denoising methods, FBP stress image

- Figure 5.20 — 3D denoising methods, FBP rest image

In the 3D case, the results have been displayed for slice # 24/32.
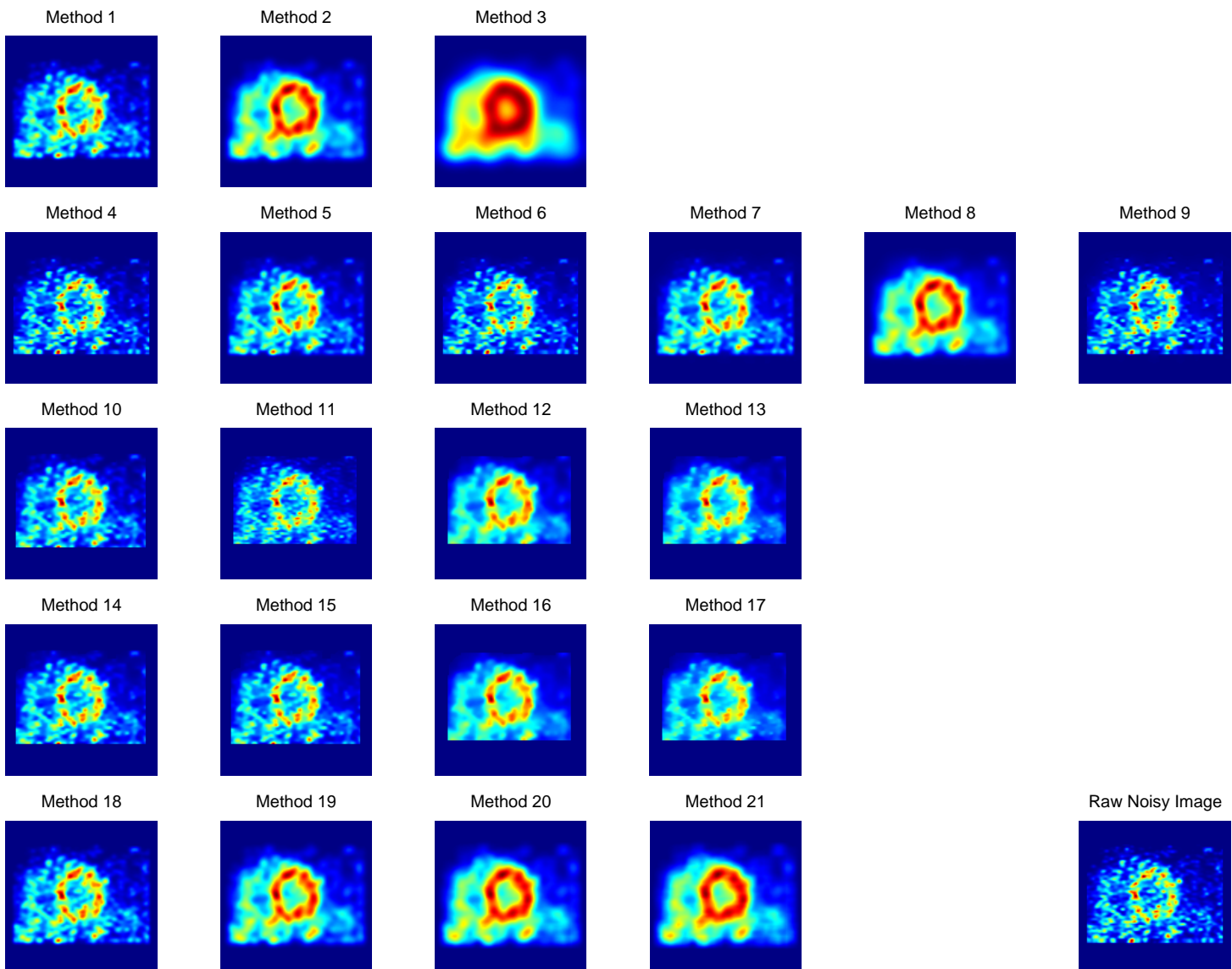
Figure 5.15: Visual results of 2D denoising protocols on clinical data, OSEM stress study
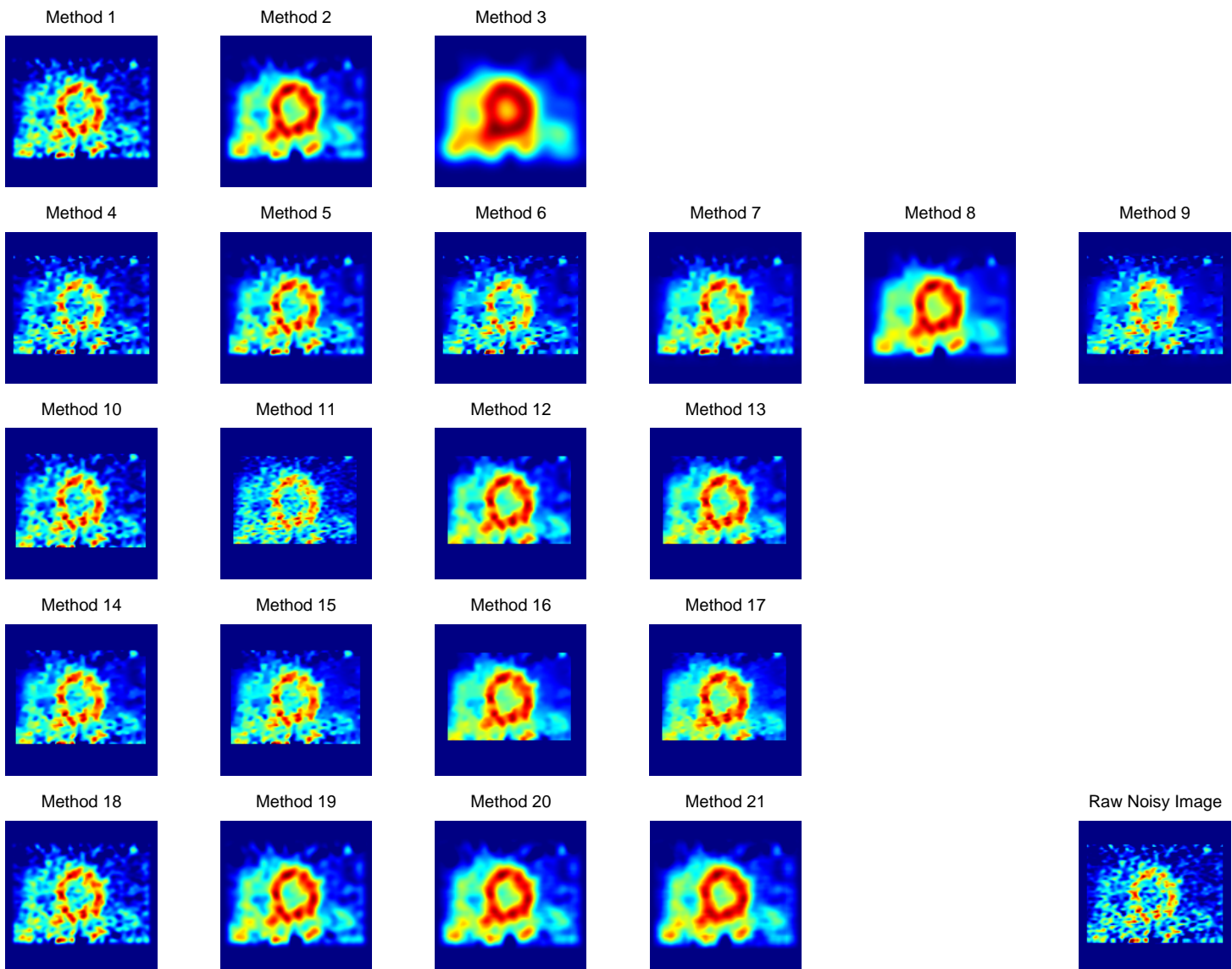
Figure 5.16: Visual results of 2D denoising protocols on clinical data, FBP stress study
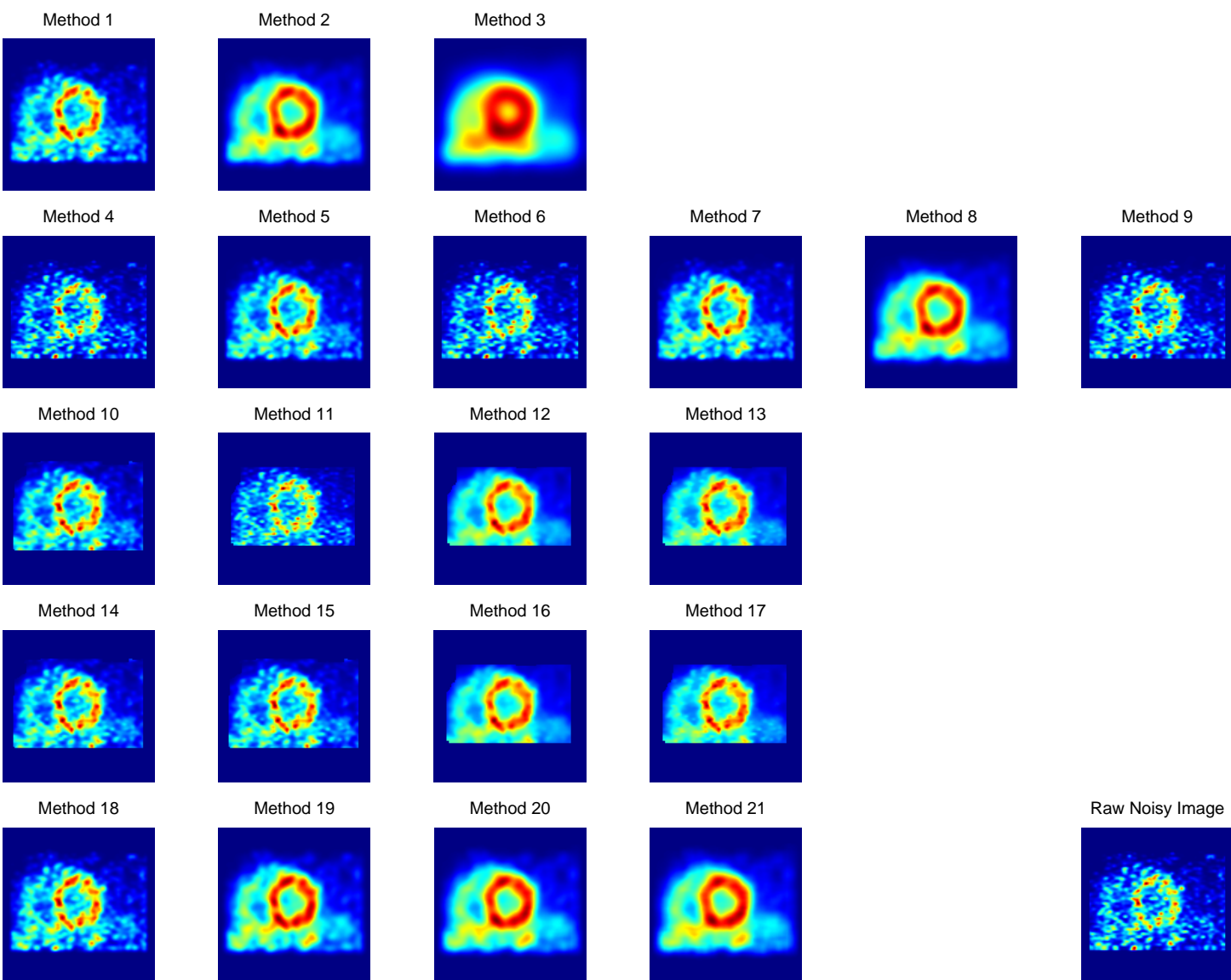
Figure 5.17: Visual results of 3D denoising protocols on clinical data, OSEM stress study, slice 24 shown
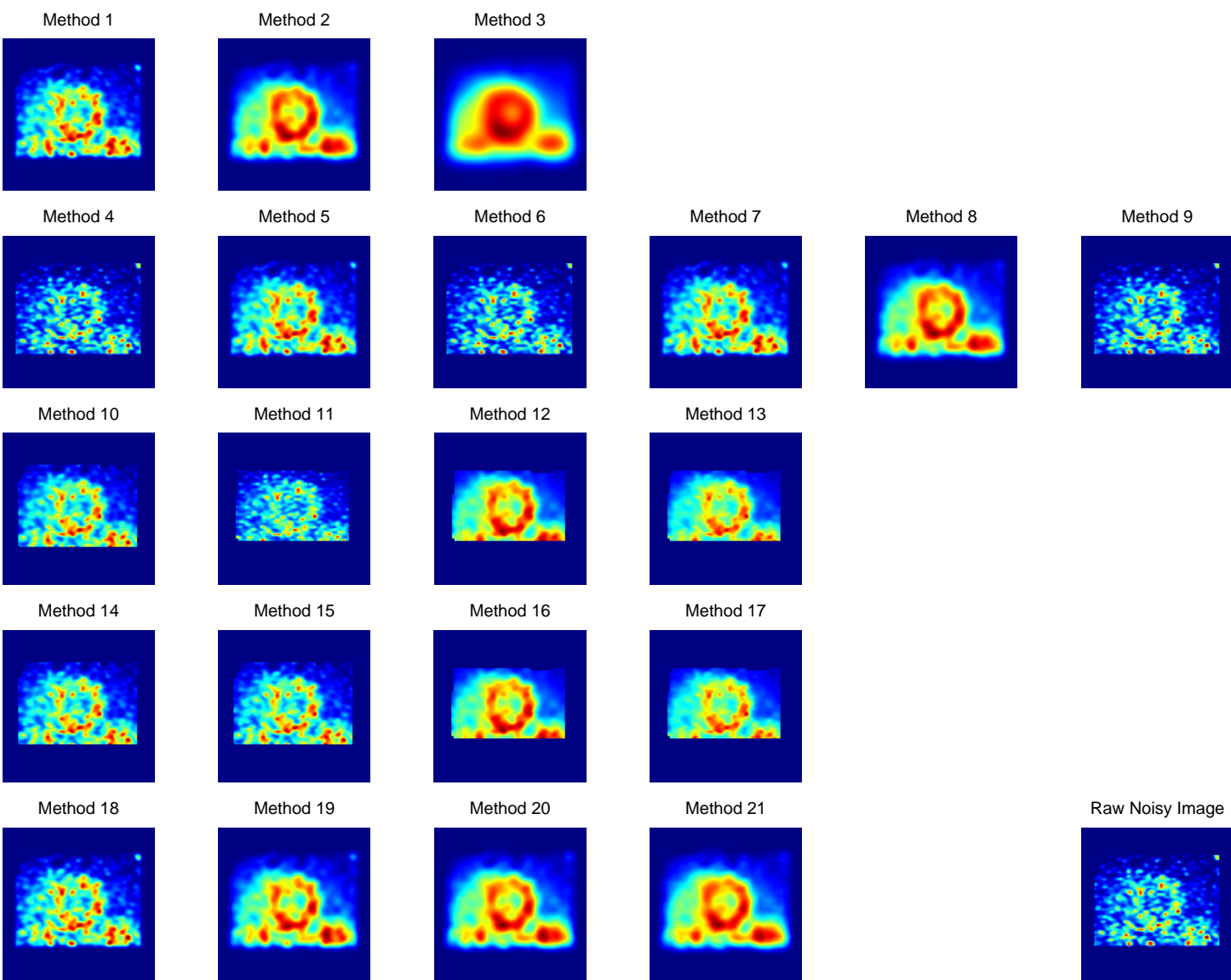
Figure 5.18: Visual results of 3D denoising protocols on clinical data, OSEM rest study, slice 24 shown
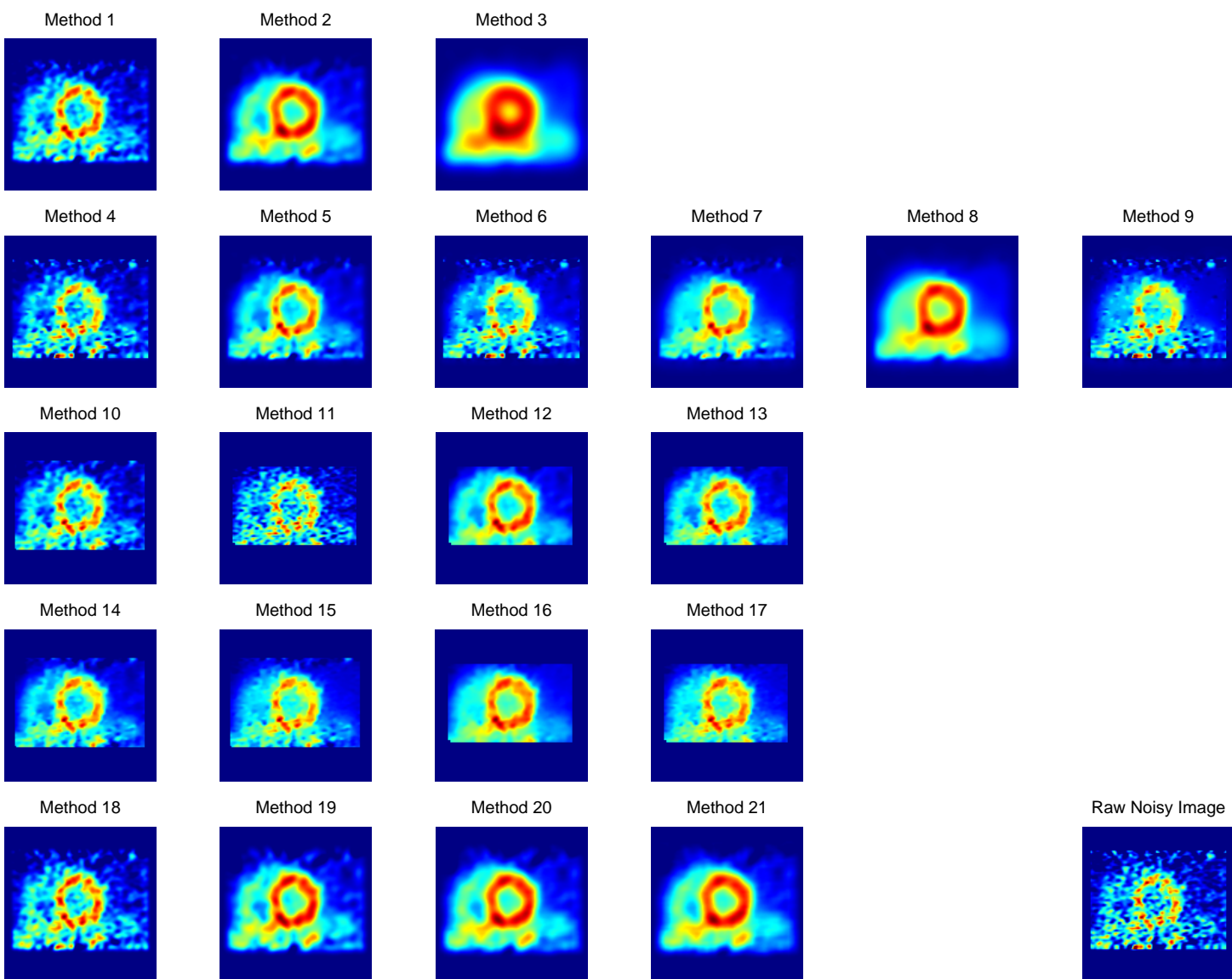
Figure 5.19: Visual results of 3D denoising protocols on clinical data, FBP stress study, slice 24 shown
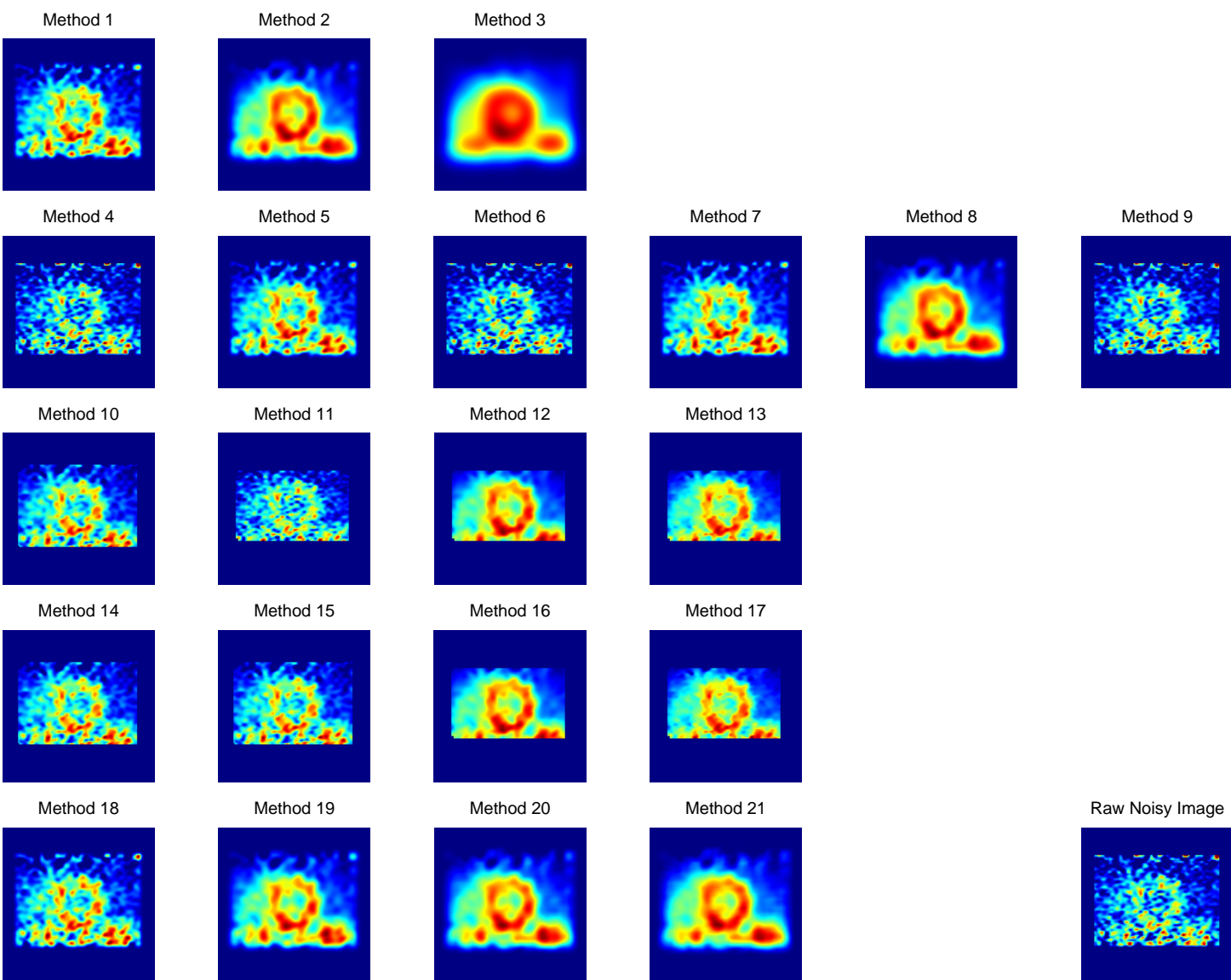
Figure 5.20: Visual results of 3D denoising protocols on clinical data, FBP rest study, slice 24 shown

Dynamic Data – Quantitative Model Output Results

Figure 5.21 plots the figures of merit extracted from the quantitative model when the denoising protocols are applied to the OSEM stress study, and Figure 5.22 plots the figures of merit extracted from the quantitative model when the denoising protocols are applied to the OSEM rest study.

This is done for both 2D and 3D denoising (FBP is not considered here). For dynamic data, we have only included *one* Gaussian-denoised result with FWHM=12mm, the value used at OHI for dynamic studies. This appears as Protocol 18.
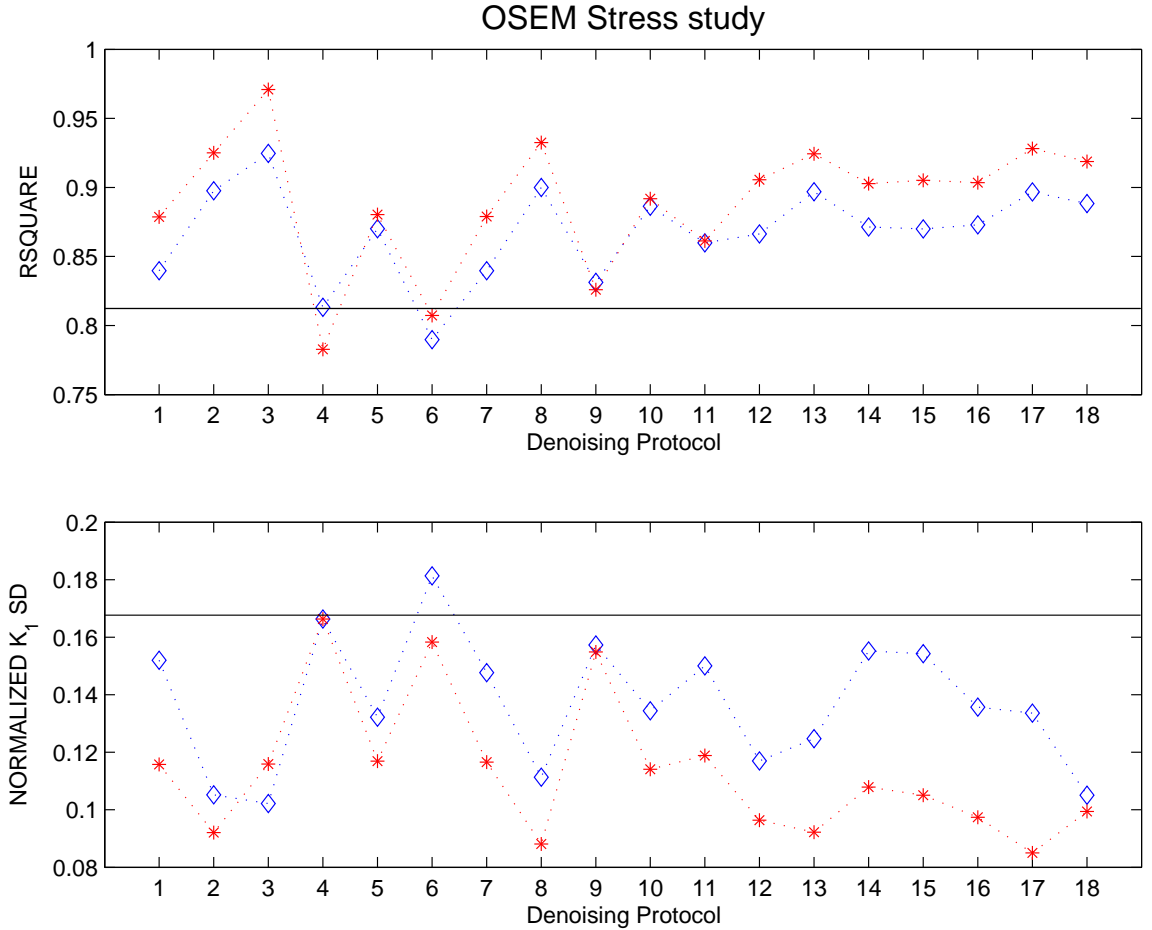
Figure 5.21: R^2 and $SD_{K1,norm}$ results for dynamic PET data (OSEM stress study). The horizontal lines denote the values derived from noisy data.
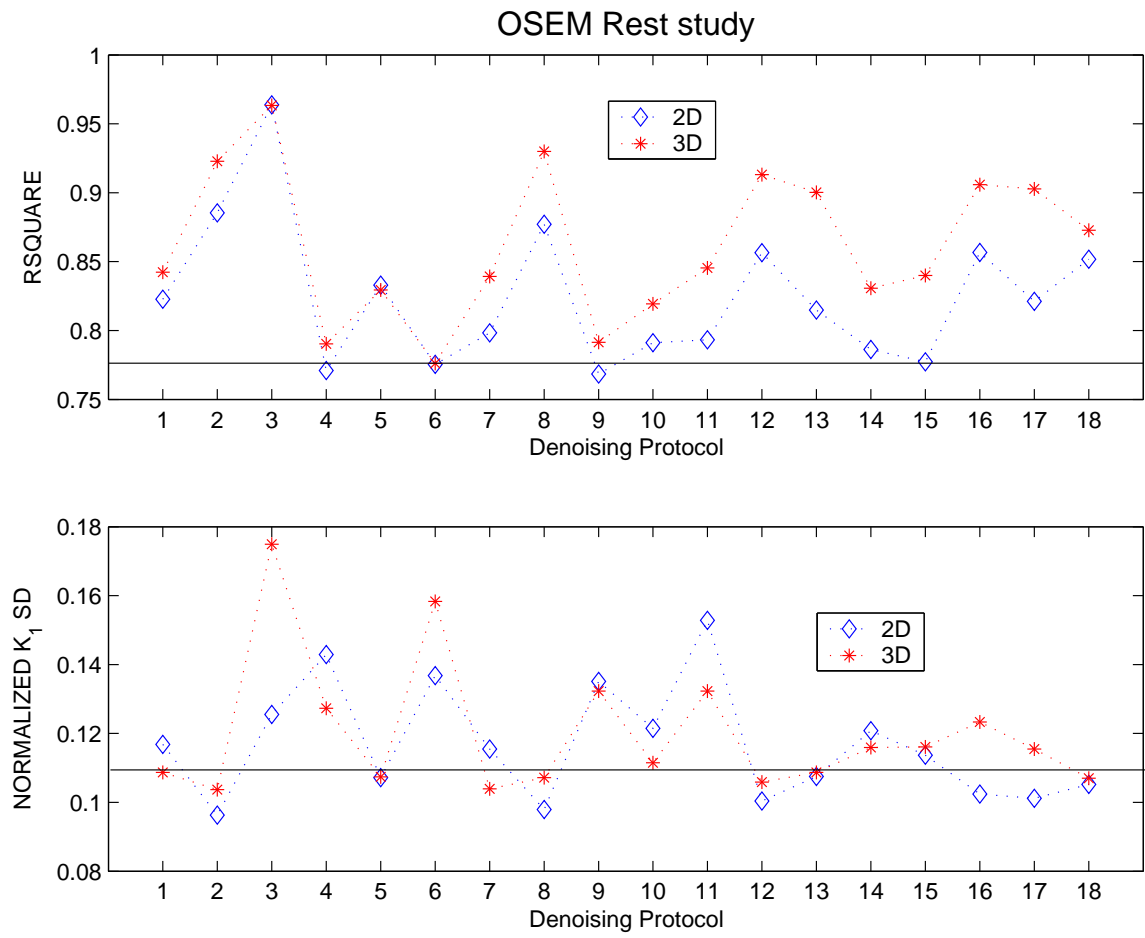
Figure 5.22: R^2 and $SD_{K1,norm}$ results for dynamic PET data (OSEMrest study).The horizontal lines denote the values derived from noisy data.

5.7 Discussion

By analyzing the experimental data and images presented in the previous section, we can draw several conclusions about the denoising protocols that were investigated. These are presented below, along with some relevant discussion points that help interpret the results of this study.

1. **Full Range of Denoising Results** – In this thesis, we have elected to include the results for all protocols, even though it is clear in some cases the denoised images are largely oversmoothed or undersmoothed. Similarly, some of the protocols generate denoised results that are essentially identical to others (*e.g.* Protocol 14 and 15. This has been done to clearly illustrate the effect that changes to both *a)* the modification methods and *b)* the modified levels has on the denoised result. By presenting a comprehensive "collage" of visual outputs, it helps the reader develop an intuition as to how these factors incrementally affect the output image quality.

2. **Classification of Denoising Protocols based on Image Quality** – By analyzing the results visually (both phantom and clinical), we notice that there is a wide range of variation in quality for a given set of input data. The denoised images can be loosely classified according to the degree of smoothing (or lack thereof) that each denoising protocol generates. Table 5.4 indicates this grouping of protocols for the set of visual results presented. Because we have established that 3D denoising gives improved quality over 2D, we restrict our attention to the 3D results. For the same reason, we consider only OSEM– reconstructed images in this analysis (while realizing that the denoising methods offer the same relative improvements for FBP reconstructed images). For the phantom images, we can easily perform this classification because we know the properties of the desired image. For the clinical images, subjective assessment

has been performed based on a set of criteria which was provided after consulting with expert PET staff at OHI, as explained in Section 5.5.2.

	PHANTOM	**CLINICAL**	
	OSEM	**OSEM** **stress**	**OSEM** **rest**
Protocols giving **undersmoothed** images	11	4,6,9,11	4,6,9,10,11,14,15
Protocols giving **marginally** **undersmoothed** images	4,6,10	1,5,7,10,14,15	1,5,7,13,17
Protocols giving **oversmoothed** images	3,8	3	3
Protocols giving **marginally** **oversmoothed** images	2	2,8	2,8
Protocols giving **appropriately** **smoothed** images	1,5,7,9,12,13, 14,15,16,17	12,13,16,17	12,16

Table 5.4: Classification of denoising protocols based on level of smoothing

3. **Denoising Protocols Giving Best Image Quality** – Table 5.4 reveals that denoising protocols 12 and 16 consistently produced denoised images that give an acceptable level of smoothing, over a wide range of input data quality. This has been verified for the FBP data separately (though the full results are not presented). Because these protocols yield excellent denoised results, it is instructive to consider these in detail before moving on.

Figure 5.23 allows a comparison of these protocols for a clinical OSEM stress study (protocols 2, 8, 13 and 17, also judged to give very good subjective image quality, are included). These images show that protocols 13 and 17 retain more detailed features around the inner and outer edge of the myocardium ring than with protocols 12 and 16. This is done without introducing excess graininess in the rest of the image (*i.e.* in the background). Indeed, this is the expected result when the cross scale regularization (CSR) technique is used. Protocols 12

and 16 use CSR, but not for the finest scale (level 1) detail coefficients, which are set to zero. Protocols 13 and 17, however, use CSR for both level 1 and 2 detail coefficients. This allows us to conclude that for this data, the benefits using CSR are fully realized only when it is applied across multiple scales. This is a different result from that found in [26], where CSR was only applied to level 1 detail coefficients. Figure 5.24 presents the outputs of the same six protocols on a clinical OSEM rest study. Note in this case that protocols 13 and 17 give a undersmoothed result, so in this case using CSR on the finest scale is not appropriate.

We conclude from the above that the value of using CSR approach lies in its use on the level 2 coefficients. In addition, the similarity of the protocol 12 and 16 results (and those of protocols 13 and 17) for both stress and rest cases indicate that spatially adaptive thresholding of the level 3 coefficients gives little additional benefit, since essentially the same result is obtained by leaving these coefficients unmodified.

4. **3D Denoising vs. 2D Denoising Image Quality** – The clinical static images and the clinical dynamic images illustrate that better denoising results are achieved by applying a 3D transform to the entire data volume, compared to "slice–by–slice" 2D denoising. For static images, the denoised image quality is generally better for the 3D denoising protocols than for the 2D versions. Specifically, 3D processing suppresses a lot of graininess in the myocardium area that appears otherwise in the 2D results. It should be noted that this behaviour was not evident in the 3D phantom studies. Comparing the raw noisy images for 2D and 3D reveals that the level of noise in the 3D phantom was significantly higher than the simulated 2D phantom noise. It is probably this increased noise level, rather than a shortcoming in the 3D denoising technique, that led to an absence of significantly increased image quality, as is evident in the clinical data
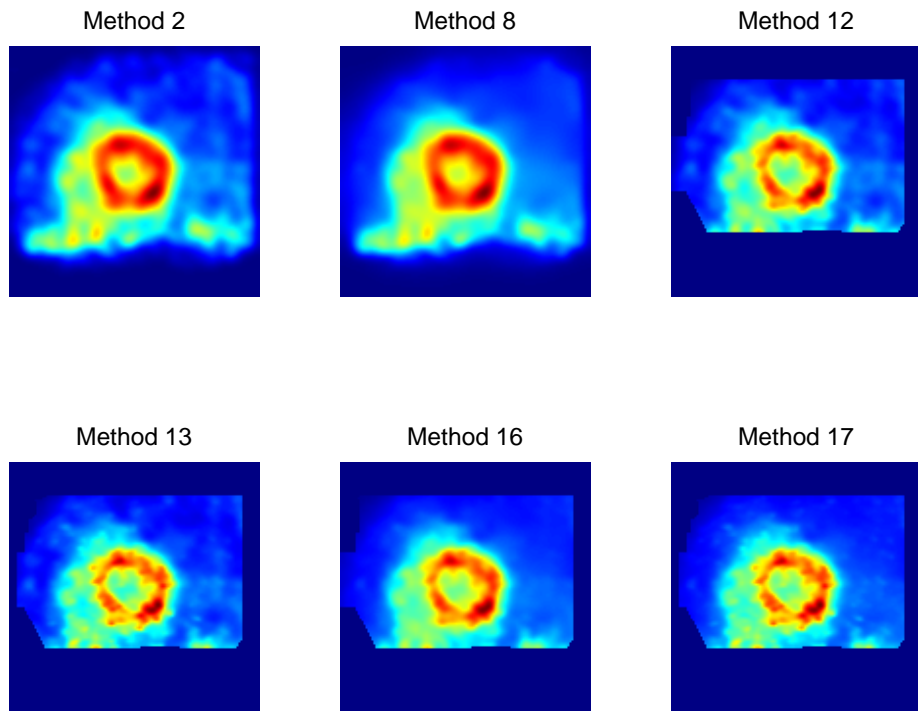
Figure 5.23: Visual results of protocols 2,8,12,13,16,17 (3D, OSEM, stress, slice 24 shown)
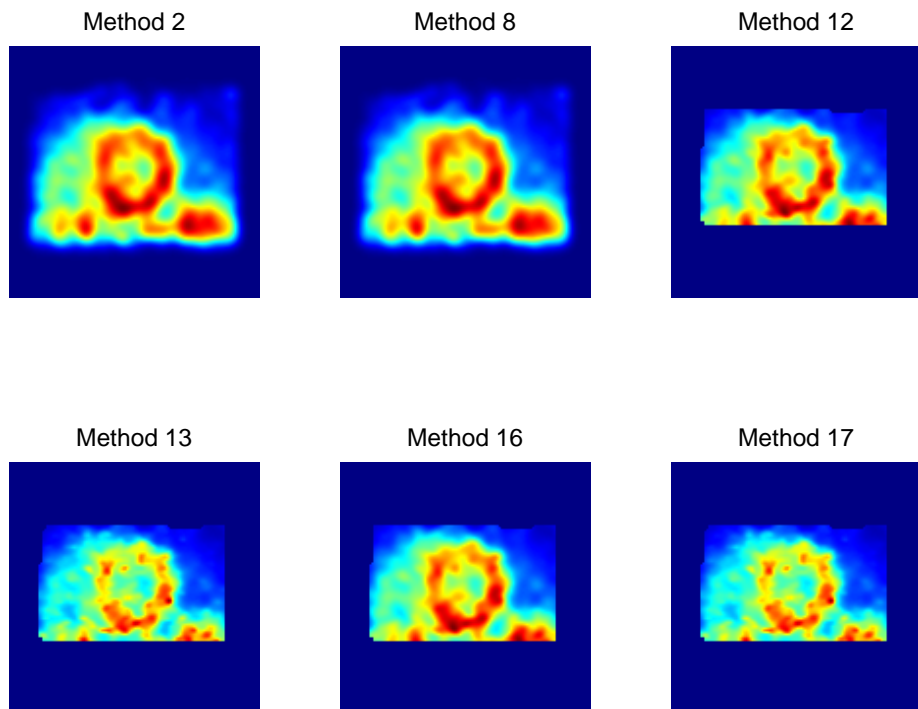

Figure 5.24: Visual results of protocols 2,8,12,13,16,17 (3D, OSEM, rest, slice 24 shown)

(discussed below).

5. **Comments on Protocols 1–3: Denoising Protocols using subband inclusion and deletion** – These protocols, very simple in that coefficient modification is done by setting entire detail subband levels to zero, provide valuable information. Protocol 3 illustrates that removing the detail coefficients at levels 1-3 from the reconstruction will give oversmoothed results, so we conclude from this that there is useful detail at these levels. Protocol 2, despite its simplicity (levels 1 and 2 set to zero, level 3 unmodified), gave surprisingly good results that were only marginally smoothed, and would certainly be acceptable in some circumstances. However, in order to get more detail we must implement a more sophisticated thresholding method on the level 1 and 2 coefficients.

6. **Comments on Protocols 4–9: Denoising Protocols using SA** – It is evident that protocols 4, 6, and 9 give very similar unacceptable results that are characterized by too much graininess. In each of these, spatially adaptive thresholding is performed on the level 1 detail coefficients, so we conclude that this method does not provide aggressive enough thresholding at this level. In protocols 5 and 7, the level 1 detail coefficients are set to zero and these show only marginal undersmoothing. The SA thresholding in protocols in 5 and 7 is done on levels 2 and 2-3 respectively, and these yield nearly identical denoised images. This suggests that applying the SA technique to level 3 coefficients provides minimal improvement in quality. Among this group, protocol 8 is judged to produce the best quality, while noting the similarity to protocol 2 above.

7. **Comments on Protocols 10–13: Denoising Protocols using CSR** – We have discussed protocols 12 and 13 above, arguing that the increased detail that protocol 13 retains around the myocardium boundary can be attributed to

applying the CSR to multiple detail levels (1-2), instead of discarding the level 1 coefficients. Protocol 11 exemplifies that although using CSR on the level 2 coefficients is beneficial, this improvement is offset by simply leaving the level 1 detail coefficients unmodified. Finally, a comparison between the results of protocols 4 and 10 shows the increase in image quality that can be achieved by using CSR on the level 1 coefficients (giving only a slightly undersmoothed image), as opposed to using SA on the level 1 coefficients (very grainy).

8. **Comments on Protocols 14–17: Denoising Protocols using SA and CSR** – Having already mentioned protocols 16 and 17 as optimal results above, we see that protocols 14 and 15 do not perform as well. This is primarily due to the fact that in these protocols, CSR is applied only at the finest scale (level 1). Having said this, it does not negate the potential benefit of using a hybrid denoising scheme that combines spatially adaptive thresholding with CSR. Protocols 14 (using CSR at level 1, SA at level 2) and 15 (using CSR at level 1, SA at level 2 and 3) give images that are only marginally undersmoothed. It may be possible to generate acceptable denoised images with these methods with changes to the parameters of the SA algorithm (*e.g.* size of the neighbourhood $\mathcal{N}$ used in the variance calculation).

9. **Comparison between Gaussian smoothing and Wavelet-based Denoising** – Using the same criteria for visual assessment, we see that the Gaussian smoothed images of FWHM = 10mm and 15mm provide images judged to be only marginally lower quality than the best wavelet-denoised results. In clinical practice at OHI, in order to correct for images of varying image quality (*i.e.* noisier than what we have considered), a conservative FWHM = 18mm is applied. Since the amount of smoothing is not adaptive to noise levels, this tends to oversmooth the data in many instances. The marginal increase in denoising

performance that is achieved with wavelet-based denoising is accompanied by the ability to adapt to image quality, *based on the input data* – a clear advantage over Gaussian smoothing.

10. **Dynamic Results** – The points above have been based on visual interpretation of the denoised results. Upon inspection of Figure 5.21 and Figure 5.22, some further comments specific to the effect of the protocols on the physiological model outputs are in order. Firstly, 3D denoising in general outperformed 2D denoising for both minimizing $SD_{K1,norm}$ and maximizing R^2. We should point out the anomaly that this was not observed with the SD_{K1} results of the rest study, as the 2D methods generally did a better job of minimizing this quantity. Concentrating on the stress study, we notice that Protocol 17 (using 3D denoising) does an excellent job for both figures of merit. This is especially encouraging, given that this protocol was deemed to be near optimal using visual assessment as well.

Since Protocol 17 has shown to improve both the visual quality and extracted model parameters, we decided to investigate this technique over a wider range of data quality. This experiment measured $SD_{K1,norm}$ and R^2 using data (OSEM stress) from four healthy subjects. The results of using Protocol 17 were compared to the standard clinical protocols used for dynamic studies at OHI – 3D Gaussian filtering with FWHM=12mm – and appear in Table 5.5. Protocol 17 introduced an average of 3.9% improvement in $SD_{K1,norm}$ and a 2.0% improvement in R^2.

Figure 5.25 shows the full quantitative model output using Protocol 17. Notice how the K_1 polar map in the top left corner shows increased uniformity when compared to the output of the same model from the original noise–contaminated data in Figure 5.26.

Subject	Protocol 17 (3D)		Gaussian (3D)	
	R^2	$SD_{K1,norm}$	R^2	$SD_{K1,norm}$
Subject 1	0.9282	0.0850	0.9188	0.0994
Subject 2	0.9747	0.0578	0.9701	0.0557
Subject 3	0.8892	0.0955	0.8398	0.0949
Subject 4	0.9656	0.0745	0.9600	0.0787

Table 5.5: Comparison of quantitative model outputs between Protocol 17 and 12mm FWHM Gaussian

11. **Caveat: Using MSE as Sole Measure of Image Quality** – Certain denoising protocols (such as 3) consistently produce denoised images that are oversmoothed and not clinically useful. Despite this, the MSE values that these protocols produce are among the lowest seen. This behaviour reiterates the need to consider **both** visual and quantitative results when determining how effective a protocol behaves. In general, Figures 5.9 and 5.10 illustrate that the best denoising methods (determined by visual assessment), such as protocols 13 and 17 *tend* to give lower MSE values (as expected), but we cannot use MSE alone as an indicator of denoising efficacy.

The next chapter concludes this thesis by summarizing the work performed, then presents some useful future directions in which this research could be taken.
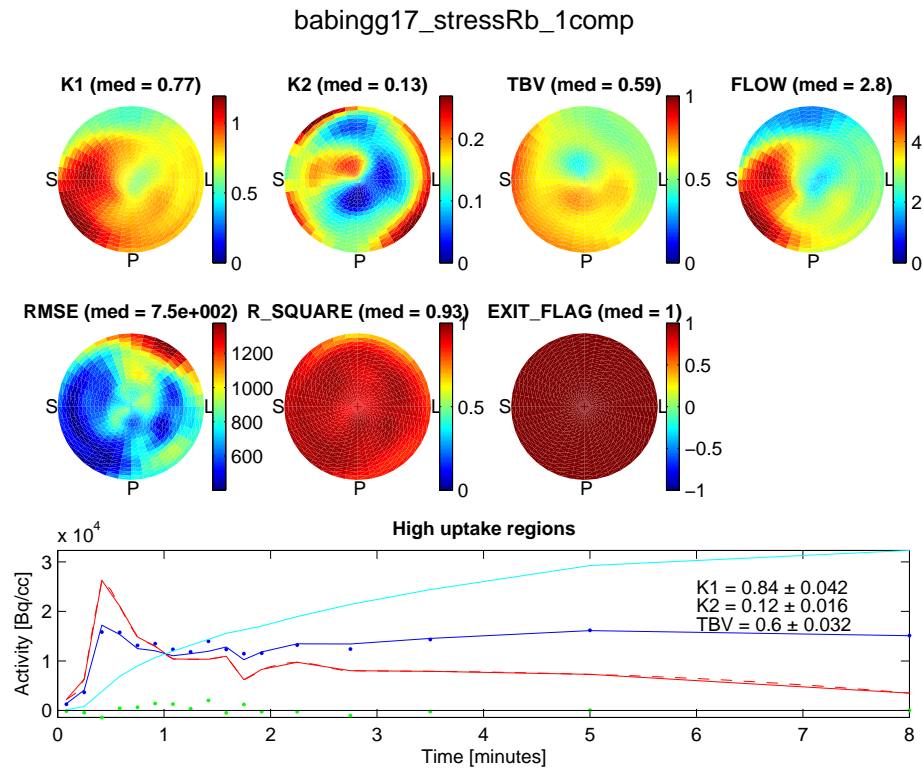
babingg17_stressRb_1comp

Figure 5.25: Quantitative model output using Protocol 17

babingg22_stressRb_1comp
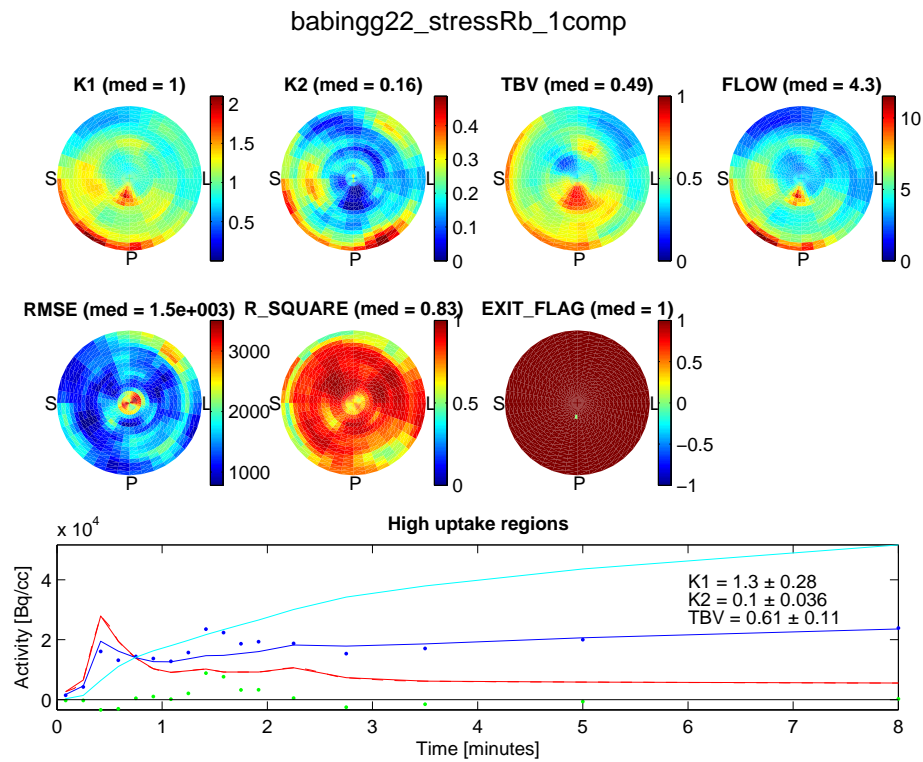
Figure 5.26: Quantitative model output using original noisy data

Chapter 6

Summary and Future Work

This final chapter summarizes this thesis and its contributions in Section 6.1. A look toward future work that might be undertaken to extend the value of this research is presented in Section 6.2.

6.1 Summary

In this thesis, we investigated the applicability of recently published wavelet–based denoising algorithms on cardiac ^{82}Rb PET data. Principles of PET physics and image reconstruction were reviewed, with particular focus on the main sources of noise in PET data. We then illustrated how PET is applied in the field of cardiology – that is, in a level of detail sufficient to be able to interpret the visual and quantitative results of the denoising protocols. A review of classical noise reduction techniques followed, and we described how the high level of noise in ^{82}Rb PET images may not be optimally suppressed with such tools (such as GLPF smoothing).

Requiring a more sophisticated approach that can adapt to varying scale and local image properties, we then turned our attention to wavelet representations. A detailed review of publications that perform denoising using wavelets was included, and we described two state of the art methods in this field – spatially adaptive thresholding

and cross scale regularization. We justified our selection of the discrete dyadic wavelet transform as our signal basis, and described an efficient filter bank implementation that can be used for processing multidimensional data.

Our interest then turned to cardiac PET data and the development of an experimental framework that could be used to investigate wavelet denoising methods. This analysis was performed on both phantom data and clinical data obtained from OHI. After describing how the noise that corrupts cardiac PET images can be characterized, estimated, and simulated, we described several figures of merit that are useful in assessing denoised image quality in the absence of *a priori* information (a difficult task faced in nuclear medicine). With these in hand, an exhaustive series of denoising protocols were developed based on the methods above. This made it possible to determine which combinations of *a)* thresholding methods, and *b)* the decomposition levels at which they were applied, made the most significant improvements to the denoised images. With the information gleaned from this investigation, we developed hybrid methods (protocols 12,13,16,17) that apply CSR at multiple levels in combination with SA thresholding. These protocols provide excellent denoising results, both visually and in terms of the quality of extracted physiological parameters. One significant advantage over standard clinical protocol is that the noise estimate is data-driven, providing an amount of smoothing that adapts appropriately to the quality of the input image.

6.2 Future Work

- **Noise Model**

  In this thesis, we developed denoising protocols based on an assumption of an additive white Gaussian noise distribution. Despite this, they exhibited satisfactory behaviour on the input PET data (that has a degree of spatial

correlation in the image domain for both FBP and OSEM, and a non–Gaussian PDF for OSEM). This was also observed in the dynamic PET study of Lin *et al.*, in which they provide a detailed discussion that explains this behaviour [3]. While these methods did lead to improvement over the current clinical protocol, it is evident from the input data (specifically, the presence of clustered residuals) that this assumption is not strictly satisfied. We feel that better denoising performance would be achieved by incorporating a more sophisticated noise model into the procedure – such models for FBP and OSEM reconstruction are developed in [41, 42]. To accomplish this, the *bootstrap method* is a promising technique by which noise properties can be measured from a single data set (*e.g.* PET scan)[50]. This method has been investigated for PET data with promising results [51, 52]. The effect of the more realistic *spatially-varying* noise parameters (σ_n and covariance matrix) can be investigated in the wavelet domain to further refine the thresholding procedure.

- **Applicability to Higher Dimensional Denoising Algorithms**

  Our denoising methods in this thesis concentrated on 2D and 3D wavelet transforms to denoise images and image volumes. Because dynamic PET data are four-dimensional (representing a time–varying volume of data), it should be possible to refine the proposed protocols to incorporate this. Specifically, a denoising algorithm based on 4D wavelet transform has the potential to perform better noise suppression than what is possible using frame–by–frame denoising. This is because such a transform would incorporate the correlation between time frames, and preliminary experiments have been performed to validate this notion. With this approach, important practical issues that must be addressed are the uniform resampling of the time dimension (which is not the case with ^{82}Rb studies that have varying frame durations), and a suitable noise estimator $\widehat{\sigma}_n$. The dynamic polar map is another representation that might potentially

benefit from these techniques.

- **Denoising in the Sinogram Domain**

  In this thesis, we performed noise reduction by post–processing images that were reconstructed with standard clinical protocol. As reviewed in Chapter 4, there have been efforts to incorporate wavelet-based denoising earlier in the reconstruction process. We believe that there is potential benefit in adapting our proposed protocols so that they can be applied to *sinograms*. By doing so, it would be possible to eliminate from the input data artifacts that result specifically from image reconstruction (*e.g.* radial streaking in FBP images). Also, statistics of the data in this domain are known to be Poisson in nature, and this may result in a more accurate noise characterization that could be incorporated into the denoising procedure.

- **Alternate Signal Basis Functions**

  In the most recent literature, there are several "non–wavelet" multiscale representations that have proved useful when applied to the image denoising problem. Among these are *platelets*, which are localized functions of varying scale, position *and orientation*. It is shown in [53] that platelets are particularly well–suited to photon limited imaging modalities such as PET, for which the Gaussian noise assumption (which is inherent to wavelet denoising) may not necessarily be true. The thresholding methods investigated in this thesis (SA thresholding and CSR) may also be beneficial when applied to the coefficients of a platelet decomposition. Other recently developed multiscale representations of similar interest include *curvelets*[54] and *brushlets*[55].

- **Comparison to Gold Standard for Accuracy of Perfusion Estimates**

  As we have emphasized throughout this thesis, the "true" spatiotemporal ^{82}Rb distribution is unknown. Though our denoising protocols have improved both

the ability to visually interpret the data and certain figures of merit for both static and dynamic studies, we have not evaluated the accuracy of our perfusion results – that is, how close they estimate the true value. This comparison can be done with the use of a *gold standard*, which for a myocardial perfusion study is performed with a $H_2^{15}O$ tracer, producing less noisy data[3]. Performing such a study would determine the extent of the bias introduced into the flow estimates with the wavelet-based methods. In a similar manner, a cardiac *phantom study* (not to be confused with the mathematical phantom we have used) would allow a comparison between a known radiotracer distribution and the output of the denoising methods.

- **Application of Denoising Protocols to other Clinical Studies**

  The denoising protocols presented in this thesis are general in nature and could potentially benefit other types of static and dynamic imaging studies, both for PET and for other imaging modalities. In particular, it would be useful to apply these methods to:

  1. ECG-gated cardiac studies, in which the image data are subdivided into separate frames within the cardiac cycle (the trigger is based on the ECG signal). These studies are useful in correcting for patient motion and allow visualization of cardiac function and determination of the LV ejection fraction.

  2. perfusion studies using different tracers, such as ^{13}N ammonia and ^{11}C acetate.

- **Optimization of Spatially Adaptive Thresholding parameters**

  When we performed SA thresholding, we used a simple equally weighted neighbourhood $\mathcal{N}$ of 3x3 (2D) or 3x3x3 (3D) coefficients for variance estimation. This gave satisfactory results, but there are some more sophisticated techniques that

appear in the literature that may prove worthwhile. Finding an optimal size of $\mathcal{N}$ and the weighting of the coefficients (based on their horizontal and vertical orientations) was explored in [25]. In [24], they include in the neighbourhood a smaller group of coefficients from the *parent subband* (the coefficients at the same spatial location, but at the next coarsest scale) in order to incorporate the spatial correlation across subbands. In addition, the iterative thresholding strategy that is used in [25] may hold promise when applied to the data in this study.

- **Significance Testing**

  This thesis has clearly shown the merits of using the proposed methods (both with visual and quantitative results) for the data sets considered, which are typical of those found in a cardiac PET setting. It is not possible to claim that these methods outperform standard clinical protocol without completing a more rigorous study. The input of expert PET staff (possibly as receiver operator characteristic study) and a larger sample size (representing a full range of image quality) would be required in order to ascertain that the improvements we demonstrate are statistically significant.

Bibliography

[1] R.E. Wahl, editor. *Principles and Practice of Positron Emission Tomography.* Lippincott, Williams and Wilking, Philadelphia, PA, USA, 2002.

[2] M. Schwaiger, S.I. Ziegler, and F.M. Bengel. *Imaging in Cardiovascular Disease,* chapter Assessment of Myocardial Blood Flow with Positron Emission Tomography, pages 195–212. Lippincott, Williams and Wilkins, Philadelphia, PA, USA, 2000.

[3] J. Lin, A. F. Laine, and S. R. Bergmann. Improving PET-based physiological quantification through methods of wavelet denoising. *IEEE Transactions on Biomedical Engineering,* 48:202–212, 2001.

[4] L. Golanowski, R.A. deKemp, R.S. Beanlands, and T.D. Ruddy. Variance and covariance of ^{82}Rb kinetic parameters: Computer simulations and dynamic PET studies. In *EMBC 2000: Proceedings of the 22th International Conference, IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA,* pages 1096–1099, 2000.

[5] J.M. Ollinger and J.A. Fessler. Positron-emission tomography. *IEEE Signal Processing Magazine,* pages 43–55, Jan 1997.

[6] CTI PET Systems, Inc. *ECAT Software Operating Instructions Version 7.2.*

[7] A. C. Kak and M. Slaney. *Principles of Computerized Tomographic Imaging.* IEEE Press, 1988.

[8] H. M. Hudson and R. S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans. Nucl. Sci.*, 13:601–609, 1994.

[9] Heart and Stroke Foundation of Canada. The Growing Burden of Heart Disease and Stroke in Canada 2003. Technical report, May 2003.

[10] D. R. Siegfried. *Anatomy and Physiology For Dummies.* John Wiley and Sons Inc., 2002.

[11] P.G. Coxson, R.H. Huesman, and L. Borland. Consequences of using a simplified kinetic model for dynamic PET data. *Journal of Nuclear Medicine*, 38:660–667, 1997.

[12] The Mathworks. *Optimization Toolbox for Matlab - User's Guide.*

[13] R.C. Gonzalez and R.E. Woods. *Digital Image Processing.* Prentice Hall, Upper Saddle River, NJ, USA, 2nd edition, 2002.

[14] S.J. Ko and Y.H. Lee. Center weighted median filters and their applications to image enhancement. *IEEE Trans. Circuits Syst.*, 38:984–993, 1991.

[15] T. Chen, K. Ma, and L. Chen. Tri-state median filter for image denoising. *IEEE Transactions on Image Processing*, 8:1834–1938, 1999.

[16] M. Unser and A. Aldroubi. A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84:626–638, 1996.

[17] M. Unser, A. Aldroubi, and A. Laine. Guest editorial: Wavelets in medical imaging. *IEEE Transactions on Medical Imaging*, 22:285–288, 2003.

[18] C. Taswell. The what, how, and why of wavelet shrinkage denoising. *Computing in Science and Engineering*, 2:12–19, 2000.

[19] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[20] D.L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inform. Theory*, 41:613–627, 1995.

[21] D.L. Donoho and I.M. Johnstone. Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association*, 90:1200–1224, 1995.

[22] B.W. Silverman. Wavelets in statistics: beyond the standard assumptions. *Phil. Trans. R. Soc. Lond.*, 357:2459–2473, 1999.

[23] R. R. Coifman and D. L. Donoho. *Wavelets and Statistics*, chapter Translation-invariant de-noising. Springer-Verlag, Berlin, Germany, 1995.

[24] S. G. Chang, B. Yu, and M. Vetterli. Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Transactions on Image Processing*, 9:1522–1531, 2000.

[25] D. Marpe, H.L. Cycon, G. Zander, and K. Barthel. Context-based denoising of images using iterative wavelet thresholding. *Proc. SPIE*, 4671:907–914, 2002.

[26] Y. Jin, P. Esser, T. Aikawa, B. Kuang, S. Duan, and A. Laine. Regularization of PET Reconstruction Using Multi-scale Adaptive Thresholding. In *EMBC 2004: Proceedings of the 26th International Conference, IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA*, pages 1616–1619, 2004.

[27] Y. Jin, E. Angelini, P. Esser, and A. Laine. De-noising SPECT/PET images using cross-scale regularization. In *Proceedings of the International Conference*

on Medical Imaging and Computer-Assisted Intervention (MICCAI), Montreal, Canada, pages 32–40, 2003.

[28] A. Stoschek and R. Hegerl. Denoising of electron tomographic reconstructions using multiscale transformations. *Journal of Structural Biology*, 120:257–265, 1997.

[29] E. Angelini, A. Laine, S. Takuma, J. Holmes, and S. Homma. LV volume quantification via spatio-temporal analysis of real-time 3D echocardiography. *IEEE Transactions on Medical Imaging*, 20:457–469, 2001.

[30] E. Kolaczyk. A wavelet shrinkage approach to tomographic image reconstruction. *J. Amer. Statist. Assoc.*, 91:1079–1090, 1996.

[31] D. Donoho. A nonlinear solution of linear inverse problem by wavelet vaguelette decomposition. *Appl. Comput. Harmon. Anal.*, 2:101–126, 1995.

[32] M. Bhatia, W.C. Karl, and A.S. Willsky. A wavelet-based method for multiscale tomographic reconstruction. *IEEE Transactions on Medical Imaging*, 15:92–101, 1996.

[33] B. Sahiner and A.E. Yagle. Reconstruction from projections under time-frequency constraints. *IEEE Transactions on Medical Imaging*, 14:193–204, 1995.

[34] B. Sahiner and A.E. Yagle. Iterative inversion of the radon transform. *IEEE Engineering in Medicine and Biology Magazine*, pages 112–117, 1996.

[35] J. Kalifa, A. Laine, and P. D. Esser. Regularization in tomographic reconstruction using thresholding estimators. *IEEE Transactions on Medical Imaging*, 22:351–359, 2003.

[36] M. Unser. Splines: a perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16:22–38, 1999.

[37] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:710–732, 1992.

[38] A. Laine, S. Schuler, J. Fan, and W. Huda. Mammographic feature enhancement by multiscale analysis. *IEEE Transactions on Medical Imaging*, 13:725–740, 1994.

[39] X. Zong, A. Laine, and E. Geiser. Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. *IEEE Transactions on Medical Imaging*, 17:532–540, 1998.

[40] I. Koren and A. Laine. *Time-Frequency and Wavelet Transforms in Biomedical Engineering*, chapter A discrete dyadic wavelet transform for multidimensional feature analysis. IEEE Press, Piscataway, NJ, USA, 1997.

[41] D.W. Wilson and B.M.W. Tsui. Noise properties of filtered-backprojection and ML-EM reconstructed emission tomographic images. *IEEE Transactions on Nuclear Science*, 40:1198–1203, 1993.

[42] E.J. Soares, C.L. Byrne, and S.J. Glick. Noise characterization of block-iterative reconstruction algorithms: I. Theory. *IEEE Transactions on Medical Imaging*, 19:261–270, 2000.

[43] E.D. Angelini, J. Kalifa, and A. Laine. Harmonic multiresolution estimators for denoising and regularization of SPECT-PET data. In *Proceedings of the 2002 IEEE International Symposium on Biomedical Imaging*, pages 697–700, 2002.

[44] S.G. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[45] The Mathworks. `http://www.mathworks.com`.

[46] The Mathworks. *Image Processing Toolbox for Matlab - User's Guide*.

[47] G. Green, A. Cuhadar, and R.A. deKemp. Spatially adaptive wavelet thresholding of rubidium-82 cardiac PET images. In *EMBC 2004: Proceedings of the 26th*

International Conference, IEEE Engineering in Medicine and Biology Society, San Francisco, CA, USA, pages 1605–1608, 2004.

[48] P. Sakellaropolous, L. Costaridou, and G. Panayiotakis. A wavelet-based spatially adaptive method for mammographic contrast enhancement. *Phys. Med. Biol.*, 48:787–803, 2003.

[49] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of gaussians in the wavelet domain. *IEEE Trans. on Image Processing*, 12:1338–1351, 2003.

[50] B. Efron. Bootstrap methods. Another look at the jackknife. *The Annals of Statistics*, 7:1–26, 1979.

[51] M. Dahlbom. Estimation of image noise in PET using the bootstrap method. *IEEE Transactions on Nuclear Science*, 49:2062–2066, 2002.

[52] I. Buvat. A non-parametric bootstrap approach for analysing the statistical properties of SPECT and PET images. *Phys. Med. Biol.*, 47:1761–1775, 2002.

[53] R.M. Willett and R.D. Nowak. Platelets: a multiscale approach for recovering edges and surfaces in photon-limited medical imaging. *IEEE Trans. on Medical Imaging*, 22:332–350, 2003.

[54] Jean-Luc Starck, Emmanuel J. Candès, and David L. Donoho. The curvelet transform for image denoising. *IEEE Transactions on Image Processing*, 11:670–684, 2002.

[55] F. Meyer and R. R. Coifman. Brushlets: A tool for directional image analysis and image compression. *Applied and computational harmonic analysis*, 4:147–187, 1997.

[56] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.

Appendix A

Background — Wavelet Theory

A.1 Continuous Wavelet Transform

The continuous wavelet transform (CWT) is, in a sense, the most "general" of wavelet representations and will be described briefly in this section. Though not of practical interest for digital images, an explanation of the CWT illustrates the mathematical concepts and provides the motivation for using the discrete dyadic wavelet transform in our denoising algorithms. This discussion will focus only on one-dimensional signals - extensions to multidimensional signals appear in subsequent sections.

The CWT is always performed with respect to a chosen continuous function $\psi(x)$, called a *mother wavelet*, and transforms a continuous signal $f(x) \in L^2(\mathbb{R})$ [13] into a function $^\psi W_C(s, \tau)$ of two continuous variables - *scale*, $s \in \mathbb{R}$, and *translation*, $\tau \in \mathbb{R}$. It is denoted with the $^\psi \mathcal{W}_C$ operator: [14]

$$^\psi W_C(s, \tau) = {^\psi \mathcal{W}_C}\{f(x)\} = \langle f(x), \psi_{s,\tau}(x) \rangle = \int_{-\infty}^{\infty} f(x) \psi_{s,\tau}(x) dx \qquad \text{(A.1)}$$

In this expression, $\psi_{s,\tau}(x)$ corresponds to versions of the wavelet function that have

---

[13] The notation $f(x) \in L^2(\mathbb{R})$ means that $f(x)$ belongs to the space of square integrable functions, *i.e.* $\int_{-\infty}^{\infty} |f(x)|^2 \, dx \leq \infty$.

[14] The notation $\langle f(x), \psi_{s,\tau}(x) \rangle$ denotes the inner product.

been shifted by the translation parameter τ, and dilated by the scale parameter s, as given by:

$$\psi_{s,\tau}(x) = \frac{1}{\sqrt{s}}\psi\left(\frac{x-\tau}{s}\right) \tag{A.2}$$

Note that $\psi(x)$ must be oscillatory and have zero average, (*i.e.* $\int_{-\infty}^{\infty}\psi(x)dx = 0$).

Some examples of $\psi_{s,\tau}(x)$ are shown in Figure A.1 for an arbitrary mother wavelet (called a Mexican hat). This figure shows that we can interpret the scale parameter s as being inversely related to frequency - as s increases, the oscillations of the wavelet function "spread out" (and the frequency of the oscillations in this local signal window decreases).
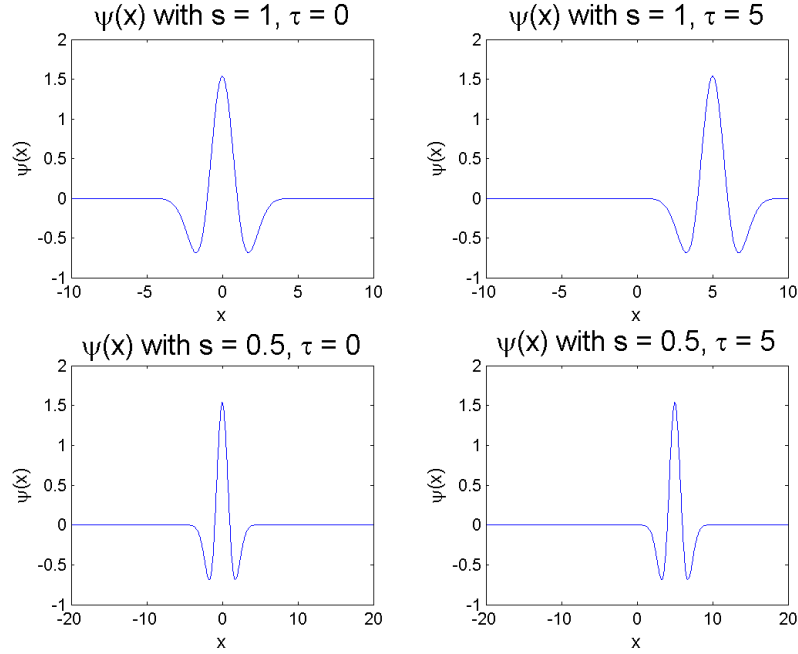

Figure A.1: Various dilations and translations of an arbitrary wavelet function

As with other signal transforms used in electrical engineering, we can interpret equation A.1 as a correlation. In this case, the CWT measures how well correlated

$f(x)$ is with shifted and dilated versions of the mother wavelet. In other words, if the signal $f(x)$ has a localized feature that matches closely with a particular version of the wavelet, say $\psi_{s_o,\tau_o}(x)$, we would expect the value of $^{\psi}W_C$ (represented as a surface in the (s, τ) plane) at the point (s_o, τ_o) to be very large. The advantage of using the CWT in locating this small feature is evident - it allows us to "localize" this information in the transform domain in a much more concise manner than in the Fourier basis, which would require an infinite number of sinusoids.

The CWT, then, can be interpreted as an infinite set of transform coefficients, $\{^{\psi}W_C(s, \tau)\}$, that measure the similarity of $f(x)$ with the infinite set of scaled, translated wavelet functions $\{\psi_{s,\tau}(x)\}$. Because both parameters s and τ are continuous, the CWT is a highly redundant transform.[15]

A.2   Discrete Wavelet Transform

Ultimately, we intend to use digital image processing techniques to denoise the PET images, so we need to eliminate the practical limitations of the CWT. This is done by using a discrete wavelet transform (DWT), as outlined in the following section, which is based primarily on the development in [13].

A.2.1   Signal Expansions as a Series

Recall that an arbitrary function $f(x) \in L^2(\mathbb{R})$ can be written as a weighted sum of expansion functions, called an *expansion series* representation:

$$f(x) = \sum_k \alpha_k e_k(x) \tag{A.3}$$

where the functions $\{e_k(x)\}_{k \in \mathbb{Z}}$ are the *expansion functions* of the series, and the scalars $\{\alpha_k\}_{k \in \mathbb{Z}}$ are the *expansion coefficients* of the series. Each value α_k gives the

---

[15]We further note that under strict admissibility conditions, the original function $f(x)$ can be reconstructed with the use of the *inverse CWT*, but the details of this are not included here.

"weight" of the signal associated with the kth expansion function $e_k(x)$.

If there is only one set of $\{\alpha_k\}$ for each $f(x)$, then the expansion is unique. In this case, the set of expansion functions $\{e_k(x)\}$ are said to form a *basis* for the class of functions being represented, and each individual $e_k(x)$ is called a *basis function*. The set of all functions that can be written in this basis (by varying the α_k coefficients) defines a *function space V*, referred to as the *closed span*:

$$V = \overline{\underset{k}{Span}\{e_k(x)\}} \tag{A.4}$$

If a function $f(x)$ is contained in that span, we write $f(x) \in V$. To compute the coefficients $\{\alpha_k\}$ for a given $f(x)$ in this space, we use the corresponding *dual functions* of the expansion set, denoted $\{\tilde{e}_k(x)\}_{k\in\mathbb{Z}}$, with the following inner product:

$$\alpha_k = \langle \tilde{e}_k(x), f(x) \rangle = \int_{-\infty}^{\infty} \tilde{e}_k(x) f(x) dx \tag{A.5}$$

There are three cases of interest:

1. The expansion set $\{e_k(x)\}$ forms an orthonormal basis for V, that is:

$$\langle e_m(x), e_n(x) \rangle = \begin{cases} 0 & \text{if} \quad m \neq n \\ 1 & \text{if} \quad m = n \end{cases}$$

   If this is the case, $e_k(x) = \tilde{e}_k(x)$, so we can compute the α_k's by substituting $e_k(x)$ for $\tilde{e}_k(x)$ in equation A.5.

2. The expansion set are an orthogonal (but not orthonormal) basis for V:

$$\langle e_m(x), e_n(x) \rangle = 0 \quad \text{for} \quad m \neq n \tag{A.6}$$

In this case, the basis functions $\{e_k(x)\}$ and their duals $\{\tilde{e}_k(x)\}$ are *biorthogonal*:

$$\langle e_m(x), \tilde{e}_n(x) \rangle = \begin{cases} 0 & \text{if} \quad m \neq n \\ 1 & \text{if} \quad m = n \end{cases}$$

and the $\{\alpha_k\}$ are computed with equation A.5.

3. The expansion set is not a basis for V, but a spanning set in which there are more than one set of $\{\alpha_k\}$ coefficients for a given $f(x)$. The expansion functions are said to form a *frame* for which the following holds [16]:

$$C_1 \left\| (f(x))^2 \right\| \leq \sum_k |\langle e_k(x), f(x) \rangle|^2 \leq C_2 \left\| (f(x))^2 \right\| \tag{A.7}$$

for $C_1 > 0$ and $C_2 < \infty$. In this case, the expansion functions and their duals are called *overcomplete* or *redundant*.

A.2.2 Multiresolution Theory

In order to describe the DWT, it is first necessary to introduce some background concepts and results from multiresolution analysis (MRA) theory. Note that the following assumes an orthogonal expansion set, a condition that will be relaxed in subsequent sections.

An MRA uses two types of functions to represent a signal $f(x) \in L^2(\mathbb{R})$ at multiple resolutions:

1. a scaling function $\phi(x)$, which is used to generate a series of *approximations* of $f(x)$, and

2. a wavelet function $\psi(x)$, which is used to generate the *details* representing the difference between successive approximations.

---

[16]The notation $\|f(x)\|$ denotes the norm, that is: $\|f(x)\| = \int_{-\infty}^{\infty} f(x)f(x)dx$

Scaling Functions

Consider a very specific expansion set, the members of which are a discrete set of scaled, dilated, and shifted versions of a function $\phi(x) \in L^2(\mathbb{R})$ called a *scaling function*:

$$\{e_k(x)\} = \{\phi_{j,k}(x)\} = \left\{2^{j/2}\phi(2^j(x-k))\right\}_{(j,k)\in\mathbb{R}^2} \tag{A.8}$$

where 2^j determines the amount by which $\phi(x)$ is compressed/stretched, $2^j k$ determines the translation of $\phi(x)$ along the x axis, and $2^{j/2}$ is a scaling factor.

Next, we see how this expansion set might form a basis for $L^2(\mathbb{R})$. Fixing $j = j_o$ gives us a subset of these expansion functions, $\{2^{j_o/2}\phi(2^{j_o}(x-k))\}_{k\in\mathbb{Z}}$, in which the functions have constant width (scale), but vary over translations. The class of functions that can be built with a linear combination of this (reduced) expansion set, defines the V_{j_o} space:

$$V_{j_o} = \overline{\underset{k}{Span}\{2^{j_o/2}\phi(2^{j_o}(x-k))\}} \tag{A.9}$$

in which a function $f(x) \in V_{j_o}$ is written:

$$f(x) = \sum_k \alpha_k \phi_{j_o,k}(x) = \sum_k \alpha_k 2^{j_o/2}\phi(2^{j_o}(x-k)) \tag{A.10}$$

Clearly, using a single j_o scale selection limits the size of the details possible in such a function. Consequently, V_{j_o} is a *subspace* within $L^2(\mathbb{R})$. As j_o increases, the expansion functions become "skinnier" and increasingly fine details can be represented in V_{j_o}.

In order to form a valid MRA, the scaling function $\phi(x)$ must satisfy the following conditions:

1. $\phi(x)$ and its integer translates $\{\phi(x-k)\}_{k\in\mathbb{Z}}$ must be mutually orthogonal,

2. the subspaces spanned by $\phi(x)$ at lower scales are nested within those spanned by $\phi(x)$ at higher scales (see Figure A.2),

3. if $f(x) \in V_j$ then $f(2x) \in V_{j+1}$,

4. the only function common to all V_j is $f(x) = 0$, and

5. as $j_o \to \infty$, any function $f(x) \in L^2(\mathbb{R})$ can be represented with arbitrary precision with the $\{2^{j_o/2}\phi(2^{j_o}(x-k))\}_{k \in \mathbb{Z}}$ expansion set.
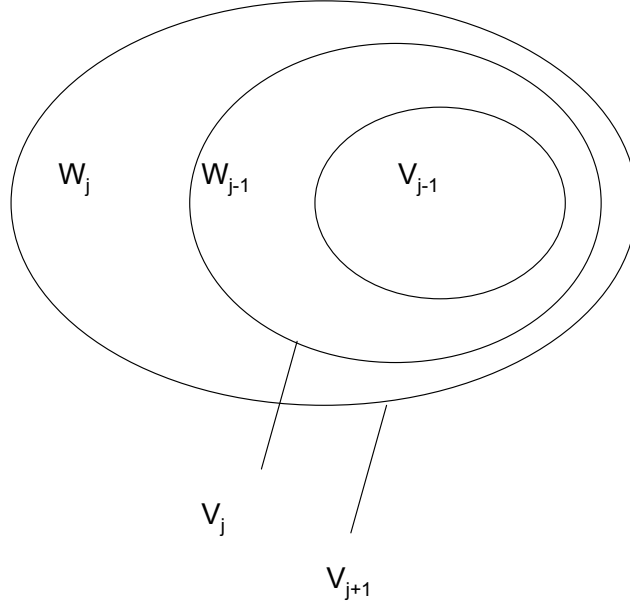

Figure A.2: a) Nesting of V subspaces, b) Nesting of V/W subspaces

If these conditions are met, the relationship between the members of the scaling function expansion sets at scales j and $j+1$ is:

$$\phi_{j,k}(x) = \sum_n \alpha_n \phi_{j+1,n}(x) \tag{A.11}$$

Substituting the original definition of $\{e_k(x)\}$ in equation A.8, replacing α_n with the more "digital filter-like" notation $h_\phi[n]$ (for reasons that will soon be evident),

and realizing that $\phi_{j,k}(x)$ evaluated at $j = 0, k = 0$ becomes $\phi_{0,0}(x) = \phi(x)$, gives the *MRA refinement equation for scaling functions*:

$$\phi(x) = \sum_n h_\phi[n]\sqrt{2}\phi(2x - n) \qquad (A.12)$$

This key result implies that the expansion functions for an arbitrary subspace V_{j_o}, can be defined in terms of the same expansion functions at the next higher subspace V_{j_o+1}. The weightings that define this relationship are given by the $h_\phi[n]$ sequence, called the *scaling function coefficients*.

Wavelet Functions

We have seen that an MRA based on scaling functions can represent a signal $f(x) \in L^2(\mathbb{R})$ with arbitrary precision, using sufficiently "skinny" scaling functions. It is important to be able to express the difference in detail between adjacent scaling subspaces V_{j_o} and V_{j_o+1}. For this reason, we introduce another expansion set based on a *wavelet function* $\psi(x)$ and its scaled, dilated and translated copies:

$$\{e_k(x)\} = \{\psi_{j,k}(x)\} = \left\{2^{j/2}\psi(2^j(x-k))\right\}_{(j,k)\in\mathbb{R}^2} \qquad (A.13)$$

where 2^j determines the amount by which $\psi(x)$ is compressed/stretched , $2^j k$ determines the translation of $\psi(x)$ along the x axis, and $2^{j/2}$ is a scaling factor.

This expansion set of wavelet functions (at a particular scale j_o) defines a function space W_{j_o}:

$$W_{j_o} = \overline{\underset{k}{Span}\{2^{j_o/2}\psi(2^{j_o}(x-k))\}} \qquad (A.14)$$

which can be used to represent $f(x) \in W_{j_o}$ as follows:

$$f(x) = \sum_k \beta_k \psi_{j_o,k}(x) = \sum_k \beta_k 2^{j_o/2}\psi(2^{j_o}(x-k)) \qquad (A.15)$$

W_{j_o} corresponds to the difference between the scaling function spaces V_{j_o} and V_{j_o+1}. That is, the set of all functions that can be represented in V_{j_o+1} but not in V_{j_o} is contained in W_{j_o}, as illustrated in Figure A.2. Thus, W_{j_o} is the *orthogonal complement* of V_{j_o} in V_{j_o+1}, and is said to obey the *direct sum* relationship:

$$V_{j_o+1} = W_{j_o} \oplus V_{j_o} \tag{A.16}$$

This relationship holds for any level j, so we can decompose V_j into the scaling and wavelet function spaces at the next lowest resolution step. If we start at an arbitrary scale j_o, and continually apply this relationship, we can represent the $L^2(\mathbb{R})$ function space in terms of scaling function spaces only, or as a combination of scaling and wavelet spaces:

$$
\begin{aligned}
L^2(\mathbb{R}) &= V_\infty \\
&= V_0 \oplus W_1 \oplus W_2 \oplus \ldots \\
&= V_{j_o} \oplus W_{j_o} \oplus W_{j_o+1} \oplus \ldots
\end{aligned}
\tag{A.17}
$$

Because the wavelet space W_j is part of a larger scaling space V_{j+1}, any function in W_j can be written in terms of the V_{j+1} expansion functions:

$$\psi(x) = \sum_n h_\psi[n]\sqrt{2}\phi(2x - n) \tag{A.18}$$

Comparing this to equation A.12, we realize that this is a *MRA refinement equation for wavelet functions*. In this case, the weightings are given by the $h_\psi[n]$ sequence, called the *wavelet function coefficients*. The scaling function coefficients $\{h_\phi(n)\}$ and the wavelet function coefficients $\{h_\psi(n)\}$ obey the following relationship:

$$h_\psi[n] = (-1)^n h_\phi[1 - n] \tag{A.19}$$

A.2.3 Representation of Signals in terms of MRA Functions

Continuous Signals

We can use the MRA "tools" developed above to represent our signal of interest, $f(x) \in L^2(\mathbb{R})$. This can be done with a *wavelet series expansion* (at an arbitrary starting scale $j = j_o$):

$$f(x) = \sum_k c_{j_o}[k]\phi_{j_o,k}(x) + \sum_{j=j_o}^{\infty} \sum_k d_j[k]\psi_{j,k}(x) \tag{A.20}$$

where $\{c_{j_o}[k]\}$ are the approximation coefficients of the series expansion (given by $\langle f(x), \phi_{j_o,k}(x)\rangle$ for an orthonormal basis) and $\{d_j[k]\}$ are the detail coefficients of the series expansion (given by $\langle f(x), \psi_{j,k}(x)\rangle$ for an orthonormal basis).

The first sum provides a lower resolution approximation of $f(x) \in V_{j_o}$. The second sum represents a set of increasingly fine details which are added to the low resolution approximation. As j increases from j_o to $j_o + 1, \ldots$, these details belong to the W_{j_o}, $W_{j_o+1}, \ldots$ subspaces. In the limit (as $j \to \infty$), enough details are available that the wavelet series becomes a precise reconstruction of $f(x)$. The similarity between the functions of equation A.20 and the corresponding subspaces of equation A.17 is evident.

Discrete Signals

Finally, we consider the case where the continuous $f(x)$ has been sampled to produce a finite discrete sequence of N samples $\{f[n], n = 0, 1, \ldots, N-1\}$. In this case, the wavelet representation of $f[n]$ is known as the *discrete wavelet transform* (DWT), which is defined at a starting scale j_o. The DWT consists of sets of *approximation coefficients* $\{{}^\phi W_D[j_o, k]\}_{k \in \mathbb{Z}}$ and *detail coefficients* $\{{}^\psi W_D[j, k]\}_{k \in \mathbb{Z}, j \geq j_o}$ measuring the correlation between the input signal $f[n]$ and the expansion functions (scaling and wavelet, respectively) at various scalings/translations. For *orthonormal* expansion

functions, the DWT coefficients are given by:

$$^{\phi}W_D[j_o, k] = \frac{1}{\sqrt{N}} \sum_n f[n]\phi_{j_o,k}[n] \tag{A.21}$$

$$^{\psi}W_D[j, k] = \frac{1}{\sqrt{N}} \sum_n f[n]\psi_{j,k}[n] \tag{A.22}$$

Note that the expressions above replace the integration of the wavelet series representation with a finite summation, reflecting the discrete nature of $f[n]$. $\phi_{j_o,k}[n]$ and $\psi_{j,k}[n]$ are discretized versions of the scaling and wavelet functions. The signal can be reconstructed using the *inverse DWT*:

$$f[n] = \frac{1}{\sqrt{N}} \sum_k {}^{\phi}W_D[j_o, k]\phi_{j_o,k}(x) + \frac{1}{\sqrt{N}} \sum_{j=j_o}^{\infty} \sum_k {}^{\psi}W_D[j, k]\psi_{j,k}(x) \tag{A.23}$$

Clearly, it is impractical to calculate these coefficients using all scales $j > j_o$, so we fix the number of scales to J. At the same time, it is useful to normalize the coarsest scale, j_o, to 0 and the number of samples, $N = 2^J$. Thus, the wavelet scales of interest are given by $j = 0, 1, \ldots, J - 1$.

Note that by considering a finite length, discrete sequence and by discretizing the scale parameter $s = 2^{-j}, j \in \mathbb{Z}$ and translation parameter $\tau = 2^{-j}k, (j, k) \in \mathbb{Z}^2$, we have overcome the practical limitations that were highlighted during our discussion of the CWT.

A.2.4   Fast Wavelet Transform (FWT)

In a manner analogous to the fast fourier transform (FFT), which has greatly increased the implementation speed of Fourier analysis, the DWT also has an efficient realization called the *fast wavelet transform* (FWT). This representation was derived by Stephane Mallat [56] and is based on the following relationships of the $\{\phi_{j,k}(x)\}$

and $\{\psi_{j,k}(x)\}$ expansion sets to the scaling function at the next highest scale:

$$\phi(2^j x - k) = \sum_m h_\phi(m - 2k)\sqrt{2}\phi(2^{j+1}x - m) \tag{A.24}$$

$$\psi(2^j x - k) = \sum_m h_\psi(m - 2k)\sqrt{2}\phi(2^{j+1}x - m) \tag{A.25}$$

A thorough derivation appears in [56], but by using these equations with the definition of the wavelet and scaling functions and the definition of the DWT, it can be shown that:

$$^\psi W_D[j, k] = \sum_m h_\psi[m - 2k]^\phi W_D[j + 1, m] \tag{A.26}$$

and

$$^\phi W_D[j, k] = \sum_m h_\phi[m - 2k]^\phi W_D[j + 1, m] \tag{A.27}$$

This provides an alternate means of computing the DWT coefficients at scale j in terms of the DWT coefficients at the next finest scale, $j+1$. Furthermore, the $\phi(x)$ and $\psi(x)$ expansion functions do not even appear in this expression. The computational advantages are gained by realizing that equations A.26 and A.27 represent convolution operations that can be implemented with a digital filter bank structure:

$$^\psi W_D[j, k] = h_\psi[-n] * {}^\phi W_D[j + 1, n]\Big|_{n=2k,k\geq0} \tag{A.28}$$

and

$$^\phi W_D[j, k] = h_\phi[-n] * {}^\phi W_D[j + 1, n]\Big|_{n=2k,k\geq0} \tag{A.29}$$

In this case, $^\phi W_D[j, k]$ and $^\psi W_D[j, k]$ are the inputs, while $h_\psi[-n]$ and $h_\phi[-n]$ are the filter impulse responses. There is a downsampling operation after the convolution. In the Fourier domain, $h_\phi[-n]$ represents a low pass filter, which generates a coarser approximation at its output, while $h_\psi[-n]$ is a high pass filter that outputs high

frequency details. Because of this frequency domain behaviour, the $\{{}^\phi W_D[j,k]\}$ and $\{{}^\psi W_D[j,k]\}$ sets of coefficients are commonly called *DWT subbands.*

Since these relationships hold at adjacent scales, this process can be repeated to create the iterated *analysis* filter bank structure shown in Figure A.3. By continuing this method L times, a set of wavelet coefficient subbands $\{{}^\psi W_D[j,k], j = J - L, \ldots, J - 1\}$ are generated. The number of times that the wavelet coefficients are calculated (L) is called the number of *decomposition levels.*

It is important to note that at the highest scale, the inputs to the overall filter bank are the DWT approximation coefficients $\{{}^\phi W_D[J,k])\}$. Assuming sufficiently fast sampling, these can be represented by the samples of the discrete input function $f[n]$.

Reconstructing a function from its DWT coefficients (*i.e.* the inverse DWT operation) can also be accomplished with a similar efficient technique. The approximation and detail coefficients at scale j_o are used to form the approximation coefficients at the next highest scale $j_o + 1$. This is repeated until we recover a signal at scale J, the original sampling scale of $f[n]$. In the *synthesis filter bank* structure shown in Figure A.3, the downsamplers are replaced with upsamplers, $h_\phi[-n]$ and $h_\psi[-n]$ are replaced with their time-reversed counterparts, and an addition operation occurs in each filter bank branch. Again, the reader is referred to [56] for a derivation.

It is not always possible to implement a FWT for an arbitrary wavelet function $\psi(x)$ (which may not have a corresponding scaling function). When it is possible, we can use the FWT to find the DWT coefficients without ever knowing what the actual MRA functions look like. This result is possible due to the extraordinary correspondence between filter coefficients and MRA functions - a wavelet basis can be defined with either. We stress that the FWT developed above pertains to an orthogonal signal decomposition - the total number of DWT coefficients (at all scales) is equal to the number of input samples.
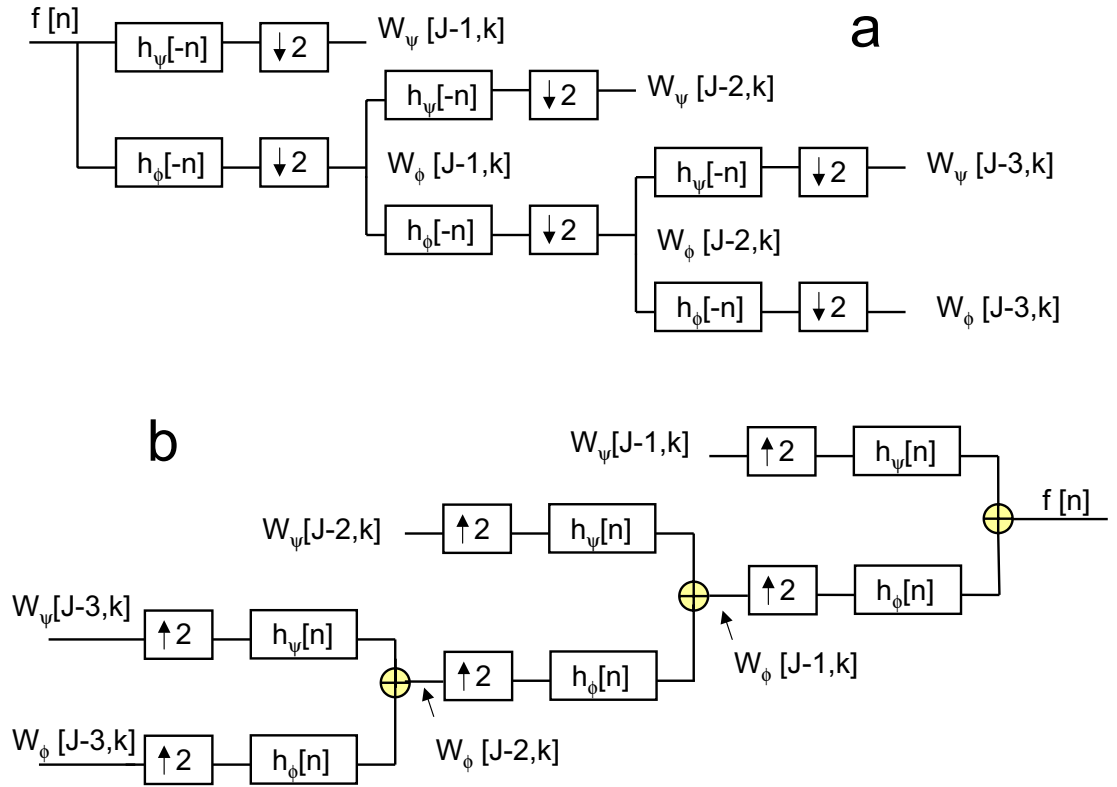
Figure A.3: A three-level FWT iterated filter bank a) Analysis b) Synthesis