DEVELOPMENT OF A MOTION CORRECTION AND PARTIAL VOLUME CORRECTION ALGORITHM FOR HIGH RESOLUTION IMAGING IN POSITRON EMISSION TOMOGRAPHY

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF MEDICAL AND HUMAN SCIENCES

2011

Shailendra Hemun Segobin
School of Medicine
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contents</td>
<td>1</td>
</tr>
<tr>
<td>List of figures</td>
<td>9</td>
</tr>
<tr>
<td>List of Tables</td>
<td>12</td>
</tr>
<tr>
<td>Abstract</td>
<td>13</td>
</tr>
<tr>
<td>Declaration</td>
<td>15</td>
</tr>
<tr>
<td>Copyright</td>
<td>17</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>19</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>21</td>
</tr>
<tr>
<td>Chapter 1</td>
<td>23</td>
</tr>
<tr>
<td>Background and Medical Context</td>
<td>23</td>
</tr>
<tr>
<td>1.0 Introduction</td>
<td>23</td>
</tr>
<tr>
<td>1.1 Alzheimer’s Disease</td>
<td>24</td>
</tr>
<tr>
<td>1.2 Imaging biomarkers</td>
<td>26</td>
</tr>
<tr>
<td>1.3 Types of imaging biomarkers</td>
<td>26</td>
</tr>
<tr>
<td>1.3.1 Structural MRI</td>
<td>27</td>
</tr>
<tr>
<td>1.3.2 PET</td>
<td>29</td>
</tr>
<tr>
<td>1.3.3 Comparison of MRI volumetry and PET</td>
<td>30</td>
</tr>
<tr>
<td>1.4 Why differing results for the hippocampus between 2 imaging biomarkers?</td>
<td>31</td>
</tr>
<tr>
<td>1.5 Motion Correction</td>
<td>32</td>
</tr>
</tbody>
</table>
1.6 Partial Volume Correction ........................................................................... 33

1.7 Conclusions .............................................................................................. 35

1.8 This thesis .................................................................................................. 36

Chapter 2 ........................................................................................................... 37

Fundamentals of PET ....................................................................................... 37

2.0 Introduction .............................................................................................. 37

2.1 Fundamental Physics in PET ..................................................................... 37

2.2 Coincidence counting in PET .................................................................... 39

2.3 Data acquisition and storage ..................................................................... 41

2.4 PET Image Reconstruction ......................................................................... 43

2.4.1 Attenuation ............................................................................................ 44

2.4.2 Scatter .................................................................................................... 46

2.4.3 Randoms ................................................................................................. 46

2.4.4 Normalisation ......................................................................................... 47

2.4.5 Calibration ............................................................................................... 47

2.5 Image reconstruction techniques ............................................................... 47

2.5.1 Analytic reconstruction methods ............................................................ 48

2.5.2 Iterative reconstruction techniques ....................................................... 49

2.5.2.1 The MLEM algorithm ................................................................. 49

2.6 The High Resolution Research Tomography (HRRT) .............................. 51
4.1.4 Extension of the observation to patient’s data ........................................ 87

4.1.5 Quantitative implications of sacrificing the accuracy of the centroid ........ 88

4.1.6 Conclusions from physical phantom experiments regarding use of centroid analysis: 91

4.2 Reframing the listmode data based on output of the centroid ....................... 92

4.2.1 Method for analysing the centroid output to reframe the listmode data .... 92

4.3 Proposed reframed-realignment algorithm .............................................. 94

4.4 Performance assessment of RAS_FBF method ........................................ 97

4.4.1 Materials for evaluation of the proposed RAS_FBF method ................. 98

4.4.2 Methods for performance evaluation of the proposed method .............. 100

4.4.2.1 Comparison of centroid and registration detected motion .............. 100

4.4.2.2 Qualitative Assessment of motion correction methods ................. 100

4.4.2.3 Assessment of motion correction methods using Time-Activity Curves 101

4.4.3 Results ........................................................................................................ 101

4.4.3.1 Comparison of centroid and registration detected motion .............. 101

4.4.3.2 Qualitative Assessment ................................................................. 103

4.4.3.3 Assessment of motion correction methods using Time-Activity Curves 104

4.5 Discussion .................................................................................................... 107

4.5.1 Impact of the work and key findings .................................................... 107
Introducing PARSLR: ........................................................................................................................... 135

A hybrid between region-based and voxel-based methods for partial volume correction in emission tomography. .............................................................................................................................. 135

6.1 Introduction ................................................................................................................................. 135
6.2 Problem Definition ...................................................................................................................... 135
6.3 Construction of the problem ...................................................................................................... 136
6.4 Efficient implementation of the LR algorithm under an EM framework ......................... 137
6.5 Reading of the LR Iterative Deconvolution Algorithm ......................................................... 137
7.6 Introducing PARSLR ................................................................................................................. 138
7.7 Reading of PARSLR ................................................................................................................ 139
6.8 Performance Assessment of PARSLR as a method for PVC .............................................. 140
6.8.1 Philosophy of assessment .................................................................................................... 140
6.8.2 Materials and Metrics used for assessment ....................................................................... 141
6.8.2.1 A Simulated hippocampus phantom .............................................................................. 141
6.8.2.2 Measured Data: 2D brain slice phantom ........................................................................ 144
6.8.2.3 Comparison with reconstruction incorporating PSF modelling ............................ 147
6.9 Results ......................................................................................................................................... 147
6.9.1 Simulated hippocampus phantom ...................................................................................... 147
6.9.1.1 Noiseless data – Complete and accurate segmentation .............................................. 148
6.9.1.2 Multiple levels of correlated Poisson noise ............................................................... 149
List of figures

Figure 1.1: Image processing steps applied to MRI and PET prior to statistical analysis between the 2 modalities. .................................................................................................................................32

Figure 1.2: Possible hypotheses as to the relationship between metabolism and atrophy ..............................................................................................................................................34

Figure 2.1: Radioactive decay example of a carbon-11 atom. ..........................................................................................38

Figure 2.2: The different types of coincidence events in PET ..........................................................................................40

Figure 2.3: Data storage methods in PET showing sinogram and listmode data storage ...41

Figure 2.4: Picture of the HRRT with the casing open ..........................................................................................52

Figure 2.5: Detector block design for the HRRT ..................................................................................................................54

Figure 2.6: Flowchart of key processes that are run via scripts on the HRRT reconstruction software ........................................................................................................................................56

Figure 3.1: Illustration of movement of an event within the brain ..................................................................................68

Figure 3.2: Illustration of LORs going in and out of the FOV as a result of motion ....................................................71

Figure 3.3: Plot of the head curve ....................................................................................................................................79

Figure 4.1: Layout of the HRRT with its 8 octagonal heads ..........................................................................................83

Figure 4.2: 2D rectilinear images ....................................................................................................................................85

Figure 4.3: Graphs showing the time course of the centroid ..........................................................................................86

Figure 4.4: Centroid data for a patient ....................................................................................................................................88

Figure 4.5: Relationship between centroid positions measured from listmode data along the Z-axis versus as measured from co-registration transformation parameters for different incidence angles ........................................................................................................................................90
Figure 4.6: Example of the time course and position of the centroid for a real patient for an [18F]-FDG scan

Figure 4.7: Flowcharts for the proposed method (RAS_FBF)

Figure 4.8: Comparison of centroid positions as measured from listmode data versus measured from transformation parameters from the registration process

Figure 4.10: Examples of motion corrected data using the specified methods with observer ratings

Figure 4.11: Time-Activity Curves of the left rectus gyrus for Subject1 and Subject2

Figure 4.12: Coronal and sagittal image of the head-rest as observed through a low resolution transmission image

Figure 5.2: Illustration of 2 possible cases of spillover effects in a middle region of interest surrounded by (i) a hot background and (ii) a cold background

Figure 6.1: 2D 4x4 pixels image (right hand side) and its corresponding B matrix (left hand side)

Figure 6.2: Left hippocampus from Hammers’ Atlas and its immediate surroundings used as mathematical phantom

Figure 6.3: Philosophies under which the PVC post-reconstruction methods of Rousset and PARSLR have been applied

Figure 6.4: Number of trues in each frame for frame1 to frame55

Figure 6.5: The digital slice of the brain selected for printing with one of the reconstructed frames shown

Figure 6.7: Graph of mean intensity of simulated hippocampus versus different Poisson random noise levels ranging from High (COV = 144.2%) to Low (COV = 4.5%) under a complete and accurate segmentation
Figure 6.8: Graph of mean intensity of Hippocampus versus different Poisson random noise levels ranging from High (COV = 144.2%) to Low (COV = 4.5%) under partial segmentation .................................................................151

Figure 6.9: Bar-chart for MSE of hippocampus across different levels of simulated Poisson noise ..................................................................................................................................152

Figure 6.10: Graph of mean activity concentration in region1 versus order of frames....153

Figure 6.11: Graph of mean activity concentration in Region2 versus order of frames...154

Figure 6.12: Rate of convergence of LR and PARSLR..............................................157

Figure 6.13: Graph of singular values for blurring kernels of different sizes presented on log scales ........................................................................................................................................161

Figure 6.14: Recovered values for Region1 with PARSLR and Rousset using different blurring kernels. ..................................................................................................................161
List of Tables

Table 1.1: Imaging Biomarkers in Alzheimer’s Disease.......................................................... 28

Table 4.1: Bed positions recorded from the scanner............................................................. 89

Table 5.1: Summary of post-reconstruction PVC methods in literature................................. 131

Table 6.1: Regions that surround the hippocampus............................................................. 142

Table 6.2: Methods of performance assessment for the 3 PVC methods............................. 143
Abstract

DEVELOPMENT OF A MOTION CORRECTION AND PARTIAL VOLUME CORRECTION ALGORITHM FOR HIGH RESOLUTION PET IMAGING

A thesis for the degree of PhD
Shailendra Hemun Segobin
The University of Manchester
September 2011

Since its inception around 1975, Positron Emission Tomography (PET) has proved to be an important tool in medical research as it allows imaging of the brain function in vivo with high sensitivity. It has been widely used in clinical dementia research with $^{18}$F-2-Deoxy-D-Glucose (FDG) and amyloid tracers as imaging biomarkers in Alzheimer’s Disease (AD). The high resolution offered by modern scanner technology has the potential to provide new insight into the interaction of structural and functional changes in AD. However, the high resolution of PET is currently limited by movement and resolution (even for high resolution dedicated brain PET scanner) which results in partial volume effects, the undersampling of activity within small structures.

A modified frame-by-frame (FBF) realignment algorithm has been developed that uses estimates of the centroid of activity within the brain to detect movement and subsequently reframe data to correct for intra-frame movement. The ability of the centroid to detect motion was assessed and the added benefit of reframe data for real clinical scans with patient motion was evaluated through comparison with existing FBF algorithms. Visual qualitative analysis on 6 FDG PET scans from 4 blinded observers demonstrated notable improvements (ANOVA with Tukey test, p<0.001) and time-activity curves were found to deliver biologically more plausible activity concentrations.

A new method for Partial Volume Correction (PVC) is also proposed, PARtially-Segmented Lucy-Richardson (PARSLR), that combines the strength of image based deconvolution approach of the Lucy-Richardson (LR) Iterative Deconvolution Algorithm with a partial segmentation of homogenous regions. Such an approach is of value where reliable segmentation is possible for part but not all of the image volume or sub-volume. Its superior performance with respect to region-based methods like Rousset or voxel-based methods like LR was successfully demonstrated via simulations and measured phantom data. The approach is of particular importance for studies with pathological abnormalities where complete and accurate segmentation across or with a sub-volume of the image volume is challenging and for regions of the brain containing heterogeneous structures which cannot be accurately segmented from co-registered images.

The developed methods have been shown to recover radioactivity concentrations from small structures in the presence of motion and limited resolution with higher accuracy when compared to existing methods. It is expected that they will contribute significantly to future PET studies where accurate quantitation in small or atrophic brain structures is essential.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright

(i) The author of this thesis (including any appendices and/or schedules to this thesis) owns any copyright in it (the “Copyright”) and s/he has given The University of Manchester the right to use such Copyright for any administrative, promotional, educational and/or teaching purposes.

(ii) Copies of this thesis, either in full or in extracts, may be made only in accordance with the regulations of the John Rylands University Library of Manchester. Details of these regulations may be obtained from the Librarian. This page must form part of any such copies made.

(iii) The ownership of any patents, designs, trademarks and any and all other intellectual property rights except for the Copyright (the “Intellectual Property Rights”) and any reproductions of copyright works, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property Rights and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property Rights and/or Reproductions.

(iv) Further information on the conditions under which disclosure, publication and exploitation of this thesis, the Copyright and any Intellectual Property Rights and/or Reproductions described in it may take place is available from the Head of School of Medicine.
Acknowledgements

My first thoughts go to my supervisors, Prof. Karl Herholz, Dr. Julian Matthews and Prof. Steve Williams. Without your help, guidance, patience and enthusiasm, I would not be submitting a PhD thesis today. Thank you will never be enough for the knowledge you imparted me with over the course of my postgraduate degree.

My postgraduate degree would not have been possible without the University of Manchester Strategic Studentship award. A big thank you to the awarding body for giving me the chance to study at a state-of-the-art imaging centre.

I would also like to acknowledge the full support I received from Prof. Geoff Parker, who, as my advisor, has always been available and whose guidance and advice helped me put my PhD in perspective.

This PhD would also not have been possible without the help and company of friends and colleagues at the WMIC. Thank you Cathleen Haense, David Coope, Georgios Angelis, Zhangjie Su, Rainer Hinz, Marie-Claude Asselin, Gerry Thompson, Matthew Jones, Matthew Walker, Stephen Carter, Jose Anton and Pawel Markiewicz for their valuable insights and the good times we spent together.

To my mum and dad without whose sacrifices, patience, upbringing and encouragement, I would never have been the person I am today, thank you will always seem like throwing a drop of water in the ocean.
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Alzheimer’s Disease</td>
</tr>
<tr>
<td>ADNI</td>
<td>Alzheimer’s Disease Neuroimaging Initiative</td>
</tr>
<tr>
<td>BGO</td>
<td>Bismuth Germanate</td>
</tr>
<tr>
<td>CT</td>
<td>Computed Tomography</td>
</tr>
<tr>
<td>FBF</td>
<td>Frame-by-frame</td>
</tr>
<tr>
<td>FBP</td>
<td>Filtered backprojection</td>
</tr>
<tr>
<td>[(^{18}\text{F})]-FDG</td>
<td>2-Fluoro-2-Deoxy-D-Glucose</td>
</tr>
<tr>
<td>FOV</td>
<td>Field of view</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full-width at half-maximum</td>
</tr>
<tr>
<td>LOR</td>
<td>Line of response</td>
</tr>
<tr>
<td>LR</td>
<td>Lucy Richardson</td>
</tr>
<tr>
<td>LSO</td>
<td>cerium-doped Lutetium oxyorthosilicate</td>
</tr>
<tr>
<td>LYSO</td>
<td>cerium-doped Lutetium-yttrium oxyorthosilicate</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>OP-OSEM</td>
<td>Ordinary Poisson-Ordered Subset Expectation Maximisation</td>
</tr>
<tr>
<td>PARSLR</td>
<td>PARtially Segmented Lucy Richardson</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
</tr>
<tr>
<td>PVC</td>
<td>Partial Volume Correction</td>
</tr>
<tr>
<td>PVE</td>
<td>Partial Volume Effects</td>
</tr>
<tr>
<td>RAS_FBF</td>
<td>Reframed Attenuation Scatter corrected_by_frame</td>
</tr>
<tr>
<td>SNR</td>
<td>Signal to Noise ratio</td>
</tr>
<tr>
<td>TAC</td>
<td>Time-Activity Curve</td>
</tr>
<tr>
<td>VOI</td>
<td>Volume of Interest</td>
</tr>
<tr>
<td>WMIC</td>
<td>Wolfson Molecular Imaging Centre</td>
</tr>
</tbody>
</table>
Chapter 1
Background and Medical Context

1.0 Introduction

In every piece of research carried out, there is the aim of understanding a process in order to make life easier, if not better. In medical research, complex medical problems are decorticated in a bit to understand a disease’s pathology in order to be able to prescribe better treatment, alleviate pain and reduce morbidity and mortality. With the rapid development and fast proliferation of medical imaging technologies, medical imaging has allowed scientists and physicians to draw-out potentially life-saving information from non-invasive investigation of the human body. The role of medical imaging has expanded from the simple visualisation and inspection of anatomic structures to becoming a tool for surgical planning and simulation, intra-operative navigation, radiotherapy planning and for tracking the progress of a disease. For example, ascertaining the detailed shape and organisation of anatomic structures enables a surgeon to preoperatively plan an optimal approach to some target structure. In radiotherapy, medical imaging allows the delivery of a lethal dose of radiation to a tumour with as little damage as possible to healthy tissue.

This work is motivated by the use and prospect of medical imaging as a biomarker for the plague that is Alzheimer’s Disease (AD) to the modern society. We also highlight the fact that accurate, repeatable and quantitative data must be efficiently extracted to support the diagnosis and follow up on the progression of AD. We limit the focus to using Positron Emission Tomography (PET) in imaging AD and to how the imaging quality is increased and quantitative measurements can be made more accurate addressing the problem of movement of patients and partial volume effects that are due to the limited resolution of emission tomographs.
1.1 Alzheimer’s Disease

Alzheimer’s Disease (AD) is a complex, slowly-progressive neurodegenerative disease. It causes gradual loss of memory, language, object recognition and even movement. It is also characterised by impaired orientation, judgement and decision-making. It was first reported in 1906 by the German psychiatrist, Dr Alois Alzheimer, in Tübingen, Germany, when his first patient, Auguste D. died following a 5-year follow-up. The first case was not only marked by the symptoms mentioned above, but also backed by microscopic images showing deposition of plaques (now known to be due to amyloid-β (Aβ) aggregation) and dense bundles of fibrils (tangles formed by the hyper-phosphorylation and aggregation of tau-protein) which are today deemed as the hallmarks of the disease.

AD today accounts for 50-60% of cases of dementia (Blennow et al. 2006). It can, occasionally, have an early onset, with the risk increasing exponentially with advancing age. It is known that 24 million people from developed regions suffered from dementia back in 2001. With an increase in life expectancy, this number is thought to double every 2 decades to reach 81 million in 2040. While supporting data from developing countries is thin, it is estimated that 60% of dementia patients live there-in. Countries most likely to suffer heavily from Alzheimer’s Disease are countries from North America, Japan and Europe, where the population is known to be ageing. Moreover, developing countries are likely to see an increase in the disease. Its increasing incidence therefore makes it of prime importance that the disease’s pathology be understood. Its early recognition should also be improved on as it is believed that delaying the onset of dementia (through drug development) by 1 year could reduce its morbidity by 9 million cases in 2050 (Hampel et al. 2010).

There are different ways of measuring the severity of the symptoms of AD. One of them is the use of clinical rating scales (e.g CDR – Clinical Dementia Rating) as well as a battery of neuropsychological tests (e.g the MMSE – Mini Mental State Examination score) to determine subjectively the severity of the disease in terms of cognitive and functional
impairment. However, the outcomes of these tests are often due to the aftermath of neurodegeneration such that the results of these tests do not have a fruitful or, more importantly, the desired impact on drug development. Thus, it does not really help in reducing the morbidity associated with the disease. In fact, at the annual Alzheimer’s Association International Conference on Alzheimer’s Disease in July 2010, it was acknowledged that the current criteria for the diagnosis of AD has never been modified since its establishment 27 years ago by the National Institute of Neurological Communicative Disorders and Stroke (NINCDS) and the Alzheimer’s Disease and Related Disorders Association (ADRSA, but now known as AA – Alzheimer’s Association). Since then, researchers have managed to understand the course of the disease better and have even been able to classify phases of a human life cycle where the disease might take different ‘turning-points’. More specifically, it is known that the first phase of amyloid deposition is pre-symptomatic and has an onset that is decades before an eventual diagnosis of dementia. The second phase, Mild Cognitive Impairment (MCI), is a stage where cognitive symptoms are present but not severe enough to impair activities of daily living. Not every patient suffering from MCI progresses to AD. By progressing into the advance stage of AD, a patient enters the phase of Alzheimer’s Dementia where he/she experiences a massive cognitive impairment that disables him from functioning independently. Working groups from National Institute on Ageing (NIA) and Alzheimer’s Association, assigned this phase of the disease (Jack et al. 2011), looking for improvement of diagnostic criteria, underlined the need for ruling out other causes of cognitive decline and for continued documentation on the progression of the decline in cognition due to AD over time. Memory impairment may not even be the main observation as issues linked to vision, reasoning, problem solving, word-finding and judgement could well be the first symptoms of AD. With measures of cognitive and functional impairment via clinical rating scales not being very accurate at the early stages of the disease, and the validity of animal models to predict efficacy in humans being limited (Blennow et al. 2006), the need for biomarkers (molecular and neuroimaging) has grown. It has therefore
been suggested that biomarker measurements be included within diagnostic criteria, despite their usefulness and reliability still being at a research level.

1.2 Imaging biomarkers

A biomarker can be defined as a measure of a normal or disease-causing biological process that can be used to analyse the potential risk posed towards developing the disease or follow-up on its progression following interventional therapy (Hampel et al. 2010). It can be derived from a body fluid (e.g., cerebrospinal fluid (CSF)) or from medical imaging. Depending on the stage of a disease, the extent of pathological changes observable in those brains vary, such that information about the spatio-temporal as well functional changes occurring within an AD brain is useful. Changes in the regional distribution of a biomarker are provided on a macroscopic scale via functional imaging modalities like Positron Emission Tomography (PET, which this thesis focuses on), Magnetic Resonance Spectroscopy (MRS, which looks into changes in the biochemical composition of tissues within the brain) or functional Magnetic Resonance Imaging (fMRI, which measures the processing of neuronal information based on the level of oxygen in the blood in the presence or absence of a stimulus) or on a mesoscopic scale through Magnetic Resonance Imaging (MRI). Detection of these spatio-temporal changes in AD brains is important as depending on the stage of the disease, the extent of pathological changes observed vary. Imaging biomarkers should be able to predict the progress of dementia in defined risk groups of patients; for example, in separating those groups of MCI patients that progress to dementia (Allegri et al. 2008).

1.3 Types of imaging biomarkers

As mentioned in the previous paragraph, there are 4 types of imaging biomarkers that have been used in AD (summarised in Table 1.1), of which MRI and PET are the most popular. This thesis focuses on the PET side of neuroimaging but the use of MRI as an imaging biomarker is briefly outlined since one major component of neuroimaging in AD
has to do with correlating results obtained from both PET and MRI (also discussed further down).

### 1.3.1 Structural MRI

Structural MRI is routinely used as a neuroimaging biomarker, not only because of its widespread availability, but also for the fact that volumetric measures taken off these high resolution images are reliable. It has been used in AD to show reduced grey matter in typical regions like the posterior association cortex, medial temporal structures like the hippocampus, amygdala and entorhinal cortex, as well as the subcortical nuclei, including the basal forebrain of the cholinergic system. In addition, these structural markers find their use in monitoring disease progression with high accuracy, with variability of manually and automatically delineated volumetric measurements observed to be less than 5% for 12 MRI scanners across a multi-centre study (Ewers et al. 2006). Over the past 2 decades, volumetric measurements have evolved from manual delineation to automatic spatial transformation, based on a template and atlas derived for that purpose, thus allowing extensions towards the use of more powerful statistical tools for correlating measurements of grey matter density, cortical thickness, white matter density and other structural derivatives. However, while being less time consuming and not requiring substantial expertise (compared to manual delineation), the optimisation of these automatic processes is still under much debate and improvement (Klein et al. 2009).

Morphometric changes within tissues or organisms do not necessarily imply a corresponding change in their function. Hence, along with structural biomarkers, it is important to obtain complementary information on the functional changes that occur within those regions and eventually establish the relationship between the functional and structural changes.
<table>
<thead>
<tr>
<th>Method</th>
<th>Measures</th>
<th>findings</th>
<th>Longitudinal studies</th>
<th>Prospects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magnetic Resonance Imaging (MRI)</td>
<td>Visual ratings of hippocampus</td>
<td>Marked difference between AD and control groups</td>
<td>Extent of atrophy progression is difficult to measure</td>
<td>-High resolution T1 imaging has already been established the hippocampus and whole brain volume as structural biomarkers and have already been successfully applied in clinical trials. -Other regions may carry more information but requires further assessment across centres.</td>
</tr>
<tr>
<td></td>
<td>Manually delineated volume of the hippocampus</td>
<td>Clear discrimination between AD vs control groups. Accuracy of 70-80% in predicting AD in MCI groups</td>
<td>Progression of atrophy at rates of 3-7% per year in AD patients versus 0.9% in control groups</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Whole brain volume analysis (automatic method)</td>
<td>Designed mainly for longitudinal studies</td>
<td>Progression of atrophy at rate of 2.5% per year in AD patients vs 0.4-0.9% in HC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patterns of cerebral atrophy in brain regions</td>
<td>Atrophy observed in cortical and subcortical areas were consistent with predictions</td>
<td>Spread of atrophy throughout the brain. Strength of relationship however difficult to establish in clinical trials</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional Magnetic Resonance Imaging (fMRI)</td>
<td>Activation studies during which blood oxygenation level depend MR signal is measured</td>
<td>Data is highly reliable within subjects between imaging sessions</td>
<td>Observed specific effects of cholinergic treatment on regional brain activation in AD.</td>
<td>Its wide availability, coupled with positive outcomes from first multicentre studies makes it a promising secondary end point.</td>
</tr>
<tr>
<td>Magnetic Resonance Spectroscopy (MRS)</td>
<td>Measurement of N-acetylaspartate (NAA) via proton spectroscopy</td>
<td>Hippocampus has reduced NAA in AD groups and in those MCI patients who progress to dementia</td>
<td>NAA increases post cholinergic treatment in AD</td>
<td>Complements information from MRI and fMRI. First multicentre study completed.</td>
</tr>
<tr>
<td>Positron Emission Tomography (PET)</td>
<td>[¹⁸F]-FDG, measures local cerebral metabolic rate of glucose (ICMRglc)</td>
<td>Reduced metabolism in parieto-temporal cortical areas with low variability in multicentre studies</td>
<td>Observed effects of cholinergic treatment on cortical metabolism in AD</td>
<td>Less available than MRI but has already proven to be a reliable biomarker in multicentre studies.</td>
</tr>
<tr>
<td></td>
<td>[¹¹C]-PIB, measures Amyloid deposition as plaques and vascular amyloid</td>
<td>Detects amyloid plaques and vascular amyloid with high sensitivity</td>
<td>Increase in PIB uptake is small with AD progression</td>
<td>-PIB useful in MCI and early AD but difficult to predict its efficiency beyond that. -Therefore better used in combination with FDG-PET. In fact, FDG and PIB complement each other well in multicentre studies. -Potential for use of cholinergic imaging as a surrogate end point is not clear yet. Microglial activation is still at a research level.</td>
</tr>
<tr>
<td></td>
<td>[¹¹C]-PK11195 Measures activated microglia in AD</td>
<td>Significant 20-35% increases in microglial activation in frontal, temporal, parietal, occipital and cingulate cortices (p&lt;0.05) of 13 AD subjects.</td>
<td>Follow up studies still in progress and yet to be reported</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[¹³C]-MP4A Measures the average reduction of Acetyl Choline Esterase (AChE) activity</td>
<td>Showed an average reduction of AChE in the brain of the order of 30-40%</td>
<td>Degree of AChE inhibition correlated with clinical improvement of cognition in AD.</td>
<td></td>
</tr>
</tbody>
</table>

*Table 1.1: Imaging Biomarkers in Alzheimer’s Disease.[updated from (Hampel et al. 2010)]*
1.3.2 PET

PET is a molecular imaging technique that allows observation of specific biological processes in vivo. It is different to functional imaging processes like fMRI where the MR signal is dependent on the blood oxygenation level (and hence the term BOLD) and therefore occurs as a result of a combination of other processes. PET, on the other hand, has a process that effectively occurs at a molecular level as is underlined with a very simplified example of $[^{18}F]$2-Fluoro-2-Deoxy-D-Glucose (FDG) PET. The tracer, FDG, has a chemical composition that is similar to glucose but has been radio-labelled with the positron emitting isotope fluorine-18 ($^{18}$F). Glucose is the main substrate for supplying energy in the brain. Like glucose, FDG is carried into tissues via the blood stream following intravenous bolus injection and undergoes phosphorylation, but no significant further metabolism. It therefore accumulates in brain tissues in proportion to the local cerebral metabolic rate of glucose (ICMRg lc) over the first 10-20 minutes post-injection and approximates ICMRg lc over the next 40 minutes. Hence, FDG PET images acquired 20 to 90 minutes post injection represent ICMRg lc in the brain. The radio-labelled $^{18}$F serves to facilitate the measurement of the distribution of the FDG from the PET scanner. Like all positron emitting isotopes, $^{18}$F has a nuclear mass that is smaller than the stable isotope of fluorine and decays via emission of a positron. The latter then travels in tissue for a short distance (1-2 mm) from the decaying nucleus before hitting an electron. The resulting annihilation produces a pair of photons (γ-rays, each 511 keV) travelling in opposite directions (the process is explained in more details in section 2.1). The PET scanner contains scintillation detectors like bismuth germinate (BGO), sodium iodide (NaI), or lutetium oxyorthosilicate (LSO) which coupled with coincident electronics detects these photon pairs and the coincident event is stored in a computer system. The line of response (LOR) is the line that connects the 2 points of photon detection and it is known that the event originated along or close to that line. With the appropriate reconstruction method, the events (several millions in a typical PET scan) are reconstructed to provide local activity concentration within the field of view (FOV) of the PET scanner. The reconstructed PET image is then analysed (both visually and
quantitatively) to determine the extent or progression of a disease. For example, in the brain, neuronal activity is reflected by IC\(MR_glc\. In AD, neuronal activity is known to be impaired and thus observed as reduced FDG uptake in the temporal and parietal association cortex, including the precuneus and posterior cingulate cortex, with the frontolateral association cortex also involved but to a lesser and variable extent (Herholz et al. 2006). The pattern seen on an FDG-PET image is reflective of the clinical symptoms of AD with noticeable decline in memory and associative thinking but with relatively preserved primary motor and sensory function that is reflected in preserved metabolism in the primary motor, somatosensory and visual cortical areas. The changes in FDG uptake can be measured more objectively and with lower coefficients of variation than standard neuropsychological measures such that changes from normal ageing to progress of the disease via response to drug therapy are monitored with more confidence (Alexander et al. 2002; Hirono et al. 2004).

### 1.3.3 Comparison of MRI volumetry and PET

MRI reproduces the structural changes happening within the brain while PET gives an insight of the functional changes. To understand the course of AD, it is desired to understand the triggers, course and chain of events that occur within the brain. More specifically, the following questions are set:

1. What occurs first: atrophy of a tissue or reduction in its metabolism?
2. With progression of the disease, is:
   a. Atrophy accompanied by reduced CMR\(glc\) in the remaining tissue?
      Showing that the progression of the disease from a structural, cognitive and functional point of view is occurring at the same rate?
   b. atrophy accompanied by sustained CMR\(glc\) in the remaining tissue?
   c. atrophy accompanied by increased CMR\(glc\)?

By combining results from different imaging modalities, researchers are able to look into these questions. However, while the patterns of hypometabolism and atrophy were in
agreement for most structures like the posterior-cingulate-precuneus, inferior-temporoparietal, parahippocampal, angular and fusiform areas (where hypometabolism exceeded atrophy, reflecting factors that induced hypometabolism like neuronal disconnections and amyloid depositions), a few hypometabolic structures, especially the hippocampus, have demonstrated different outcomes from different state-of-the-art clinical research centres. While some have shown that the hippocampus succumbs to the same fate as the above mentioned regions (Mosconi et al. 2009), others have showed relatively preserved metabolism with respect to atrophy that is interpreted as compensatory mechanisms being triggered to palliate for the disconnection processes already occurring based on the disease’s inherent pathology (Samuraki et al. 2007; Chetelat et al. 2008).

1.4 Why differing results for the hippocampus between 2 imaging biomarkers?

Figure 1.1 shows a flowchart summarising steps in image processing that need to be applied to the reconstructed MRI and PET data prior to comparing the results from the 2 modalities. It is first highly recommended that a study specific template be generated. A template that is specific to AD patients will be morphometrically different to one which is used for healthy controls for example, as AD patients are likely to have thinner cortices and enlarged ventricles. This increases the accuracy of the normalisation process as the transformation is more detailed (Good et al. 2001; Ashburner 2007). Consequently, the segmentation of the brain is also more accurate as this process is usually done in the same stride in popular segmentation algorithms like those used in SPM8.
**Figure 1.1:** Image processing steps applied to MRI and PET prior to statistical analysis between the 2 modalities. Framed in red are the processes of motion and partial volume correction that are either not applied to the PET data or not optimised.

While errors linked to image normalisation have been looked into (and still considered a challenge), efforts to bring qualitative improvements to the PET data has been sub-optimal. Framed in red in figure 1.1 are 2 correction processes that need to be applied to data on the PET front: motion correction and partial volume correction.

### 1.5 Motion Correction

Over the years, improvements in scanner resolution and *de facto*, the increasing ability to measure small changes occurring within the brain makes it that head movement is becoming a growing concern in brain PET imaging. Scanning durations in clinical research are typically long and because of movements, measured activity concentration within a small region like the hippocampus may be ‘artefactually’ altered such that effective changes in its measurement may go undetected or erroneously modified. In order to
derive full benefits of a higher resolution scanner, the problem of motion detection and correction needs to be addressed.

1.6 Partial Volume Correction

Partial Volume Effects (PVE) in emission tomography are due to the inherent limits of spatial resolution of PET cameras. They are specifically of concern when the cross-sectional distance of a structure (like the hippocampus) is less than twice the full width half maximum (FWHM) of the scanner’s point spread function (PSF). PVEs manifest themselves as a blurring of the region of interest as the region is under-sampled, with activity in the tails of the region of interest being lost. It also leads to spillover effects between neighbouring hot and cold regions causing reciprocal changes in their intensity measurements. In the context of Alzheimer’s disease, it is desirable to know whether the loss of activity within an atrophic region actually represents true loss in tissue function as opposed to loss in activity due to PVE. Figure 1.2 illustrates the problem of partial volume in small regions like the hippocampus.

However, in studies that compare atrophy from MRI with hypometabolism observed on the PET, it is observed that few research centres apply either or both correction techniques to their data prior to statistical comparisons. (Mosconi et al. 2009) do not apply any motion or partial volume correction to their longitudinal studies. In fact, most research groups do not perform any motion correction on their data and consider the application of partial volume correction algorithms only.
There are several reasons as to why motion correction is not applied. One of them is that the resolution of the PET images is low enough that the component of motion is not severe enough to bring substantial changes to quantitative measurements. However, modern age PET scanners like the ECAT-HR+ scanner has an isotropic resolution of 4.2mm at FWHM such that in a well-structured study like that of (Chetelat et al. 2008), the motion component deserves to be addressed. One potential reason as to why motion correction is not considered is that retrospective correction on data that has already been acquired is impossible. One of the reasons for motion artefacts is emission-transmission mismatch, that is, a misalignment between the emission data and the $\mu$-map, which is a
map of linear attenuation coefficients obtained by performing a *transmission scan* (the concept of reconstructing emission PET data using a transmission scan is discussed in more details in Chapter 2). Since the emission data cannot be reconstructed again, the emission-transmission mismatch cannot be accounted and corrected for. Last but not the least, there is also the absence of an accurate and reliable external instrumentation system for motion detection and measurement that is well tolerated by patients, especially in dementia studies.

Methods for partial volume correction that can be applied post-reconstruction are widely available in literature. Most of them are however based on assumption of a finite and accurate number of homogeneous regions which is not always true, especially for the region of the medial temporal lobe where the hippocampus lies close to the lateral ventricles which are known to be difficult to segment. Those lateral ventricles also house the choroid plexus which contains a lot of blood vessels. In the absence of a complete and accurate segmentation, those PVC algorithms tend to give results that are effectively biased.

### 1.7 Conclusions

A brief overview of Alzheimer’s Disease has been given in an attempt to underline its impact in today’s world and justify why understanding the disease’s progression and underlying pathology is important (for eventual drug development). The important role that imaging biomarkers are playing or are expected to play have been highlighted. One of the challenges towards achieving this is to make sure that the measured data is of high quality and accuracy.

Two of the challenges that PET data face are those of motion correction and partial volume correction. The purpose of this thesis is therefore to describe the appropriate structures for both motion and partial volume correction in clinical research. Novel methods for motion correction and partial volume correction are developed and their
performances assessed with respect to existing methods in literature, addressing the reasons why those current algorithms are not optimal.

1.8 This thesis

Following a brief venture into the clinical problem that needs to be addressed and having narrowed it down to the methodological part on the PET side where improvement is still desired, this thesis moves to giving an overview of the fundamentals of PET imaging: that is, from positron-electron annihilation to how an image is effectively formed, the inherent data correction methods that need to be accounted for during reconstruction (chapter 2). The High Resolution Research Tomograph is introduced as well as the characteristics that make its use for brain imaging attractive. The pronounced problem of motion is again underlined together with why Partial Volume Correction (PVC) is still desired. The problems of movement correction and PVC in PET are reviewed in details in chapters 3 and 5 respectively. Novel methods for motion correction (chapter 4) and PVC (chapter 6) are presented and their performances compared to existing methods in literature assessed. The thesis ends with a summary (chapter 7) of what has been achieved so far and how this sets the stage for helping in answering the clinical question of the atrophy-hypometabolism profile of the hippocampus accurately.
Chapter 2
Fundamentals of PET

2.0 Introduction

This chapter briefly explains how the spatial distribution of activity within the brain is measured using PET. The essential steps in data corrections during reconstruction, including attenuation and scatter are introduced to give an insight into why they are key for consideration during motion correction. The HRRT is also introduced and its high resolution explained. Finally, due to its developmental nature, the process of image reconstruction on the HRRT via scripts that can be adapted to enable motion correction is highlighted.

2.1 Fundamental Physics in PET

Positron emitting isotopes usually decay as according to equation 2.1, where the proton in the nucleus decays to a neutron, emitting a positron ($\beta^+$) and a neutrino ($\nu_e$).

\[ ^{\Delta A}X \rightarrow z^{\Delta A} + \beta^+ + \nu_e \]  

Equation 2.1

These positrons ($\beta^+$) travel through the surrounding tissues where they lose their kinetic energy via collisions and scattering events. The positrons mostly hit electrons undergoing Coulomb interaction and since the rest mass of the positron is the same as that of the electron, the positrons may undergo large deviations leading to a tortuous path through the body tissues as they lose their kinetic energy. When these positrons reach thermal equilibrium, they interact with electrons principally via annihilation (99.5%), producing two 511 keV photons as per equation 2.2 and illustrated in Figure 2.1

\[ \beta^+ + e^- \rightarrow \gamma + \gamma \]  

Equation 2.2
The positron-electron interaction can also lead to the formation of a positronium that in-turn exists in 2 forms in its ground state:

(i) An ortho-positronium where the spins of the electron and the positron are parallel. The ortho-positronium undergoes self-annihilation to emit 3 photons (Evans 1955)

(ii) A para-positronium where the spins are anti-parallel. It further decays via self-annihilation yielding two anti-parallel 511 keV photons.

![Diagram](image)

**Figure 2.1:** Radioactive decay example of a carbon-11 atom. The proton decays to a neutron releasing a neutrino and a positron. The latter travels along a tortuous path and eventually annihilates with an electron, releasing a pair of photons, each of 511 keV. The photons travel back-to-back in almost opposite directions.

The pair of annihilation photons produced from equation 2.2 normally travels in opposite directions as the positron has slowed down, with momentum resulting in a slight offset in co-linearity. If these annihilation photons travel through the body without further interactions, they can both be detected ‘in sync’ by the scintillation detectors of the PET scanner and thereby associated with each other to form a *coincidence event*. If the photons are collinear, then the positron must have been emitted somewhere along the
line that connects these two detectors with this line known as the line of response (LOR). The coincidence events are detected by the instrumentation systems and coincidence circuitry within the PET scanners. A PET scanner contains scintillation detectors that are made from materials like Sodium Iodide (NAI) or Bismuth Germanate (BGO), with the more modern PET scanners having detectors made from crystals like Gadolinium Oxyorthosilicate (GSO), Cerium-doped Lutetium Oxyorthosilicate (LSO) and Cerium-doped Lutetium-yttrium Oxyorthosilicate (LYSO). When 511 keV annihilation photons are incident on the crystals, the energy of the photons is absorbed and re-emitted as scintillation light which is subsequently detected by photomultiplier tubes (PMTs). The annihilation photon, incident on the scintillator, interacts either via photoelectric absorption or Compton scatter, triggering a cascade of scintillation photons that are incident on the photocathode of the PMT. The scintillation photons interact with the photocathode via the photoelectric effect, ejecting photo-electrons into the PMT (which is in vacuum). The photo-electrons are accelerated and multiplied to ultimately produce a short electrical pulse. The signal is further amplified and processed via the coincidence circuitry to locate and store the coincidence events. A schematic of HRRT PET scanner with the dual LSO/LYSO layers and connection to the PMTs are given in figure 2.4 and figure 2.5 respectively.

### 2.2 Coincidence counting in PET

The detector units of a PET scanner register coincident events, often called ‘prompts’ if they satisfy the following criteria:

1. The photon pairs are detected within a predefined time window, also called the coincidence window.
2. The subsequent LOR that originates from these photon pairs are within an acknowledged acceptance angle for the tomograph.
3. The energy emitted by the photon pairs is within the defined energy window. (Bailey et al. 2004)

Yet, all registered prompts may not be wanted as photons can get scattered or as detection of these coincident events is due to the accidental detection of positron
annihilations. Prompts therefore include true, scatter and random events. Figure 2.2 illustrates the common types of coincidence events in PET. A true coincidence event (Figure 2.2a) is one which both annihilation photons pass through the subject undeflected and are both detected. Figure 2.2b shows a scattered event—a detected event where either or both of the photons have undergone Compton Scatter at least once, either within the subject or against any matter that is within or outside the scanner’s FOV. A scattered coincidence event results in the registration of an LOR that does not pass on or close to the point of positron-electron annihilation. There is also the case of lost true coincident events, again as a result of Compton Scatter. It is effectively the other side of a scattered coincidence event as the latter might have occurred within the subject within the FOV and the LOR has gone undetected and is illustrated in figure 2.2c. Finally, a random (or accidental) coincidence event occurs by chance when two nuclei decay at approximately the same time, resulting in the emission of four photons. Two of these photons from different annihilation sites then happen to be detected within the time window and are believed to have come from one annihilation site as the other two photons are lost. The resulting LOR is spatially incorrect, as shown in figure 2.2d. These effects need to be accounted for when reconstructing the distribution of the tracer within the FOV into an image.

Figure 2.2: The different types of coincidence events in PET, and the effect of attenuation. (a) True coincidence event registered along an LOR that passes through the point of annihilation. (b) Scattered coincidence, where the LOR does not pass near the point of annihilation. (c) Attenuation (loss of a true coincidence event), instead of obtaining a true coincidence as in (a), no coincidence was detected due to attenuation as the photon passing though the scintillator is windowed out with energy discrimination. (d) Random coincidence occurring when two nuclei decay simultaneously, producing an LOR that contains neither annihilation points.
2.3 Data acquisition and storage

The coincidence events registered by the PET scanner can be generally stored in 2 formats: listmode or histogram mode as illustrated in Figure 2.3.

![Diagram of PET data acquisition and storage]

**Figure 2.3:** Data storage methods in PET showing sinogram and listmode data storage. Illustration of a point source within an object and example of one projection over angle ϕ along plane where the transverse distance is represented by S. Also presented is the sinogram of the point source for all projected planes used and highlighted in grey are the 'coordinates' of the single projection example. [modified from Nuclear Medicine lectures of Dr. Adam Alessio, University of Washington 2007]

In histogram mode, each coincident event is placed into a sinogram, with the histogramming often done on-the-fly by dedicated computer hardware and software. In histogram mode, the length of an acquired frame must be defined a priori as the sinograms are only temporarily held in the computer’s memory waiting for the histogramming process. Longer concatenated duration frames can be constructed by adding post-reconstruction, but other frame durations are not possible.

In listmode acquisition on the other hand, each coincident event is sequentially stored to a data file. Information about the type of event (whether prompt, delayed coincidence, non-coincidence like timing or gating information) is typically encoded. For a coincidence event, the detector position that recorded the photon and the corresponding LOR is also
encoded. A listmode file can also record the energy deposited by each photon and their relative detection times which could be useful in scatter correction or in time-of-flight PET.

Post-acquisition, the listmode data is processed by splitting into smaller and as many listmode files desired based on selected frame definitions. These split listmode files are then histogrammed into sinograms by starting with empty sinograms and encoding the LORs on an event-by-event basis by incrementing the corresponding sinogram bin if the coincidence event is a true coincidence or decrementing if it is a random one.

Since the data is histogrammed post acquisition, listmode data acquisition does not require substantial memory as histogram mode acquisition does. However, it does require a large amount of disk space, with the size of the listmode file dependent on the count-rate and duration of the data acquisition. Typical listmode files from PET scans acquired on the HRRT are between 10GB and 30GB depending on the tracer being used, their physical and biological half-life as well as the amount of activity uptake in the brain. However, nowadays disk space is much cheaper and with an increase in the number of detectors and the dominance of data acquisition in 3D, as well as the need for dynamic acquisitions, listmode data acquisition and storage is replacing sinograms in modern day scanners.

Listmode data acquisition is also becoming more and more popular in clinical research as they allow for dynamic PET reconstructions post scanning. This is in turn owing to the fact that timing information is retained in the listmode data, hence offering that temporal flexibility for reconstruction post acquisition (Nichols et al. 2002). More importantly, this feature is key to improving methods for motion correction which has become mandatory for PET imaging at high resolution (Bloomfield et al. 2003).

Storing histogrammed data in the form of sinograms is ideal in a clinical setting as it accelerates the whole image reconstruction process. Before the advent of iterative reconstruction algorithms, analytic reconstruction was the method of choice for reconstruction and algorithms like Filtered Back-Projection (FBP) require those projection
data for processing. The total processing time is reduced which is beneficial in cases where an image is required prior to neurosurgery in oncology for example.

For a more detailed and technical overview of PET data acquisition, the author suggests the references of (Defrise and Kinahan 1998 and Fahey 2002).

2.4 PET Image Reconstruction

In any imaging problem, the object has some quantifiable properties which allows for instrumentation systems to detect and measure them. However, the relationship between those measurements and the true value within the object is not always a straightforward conversion. Image reconstruction is the detection of these measurements and solving the resulting inverse problem taking into account factors that might have altered the measurement to eventually retrieve the true measurement of the object, or at least, the closest estimate to the truth.

In PET imaging, the object is effectively a distribution of radioactivity within the subject’s body. If this activity distribution is represented as a 3D volume consisting of $N$ number of voxels, the true image $f$ can be represented as a vector of $N$ elements where the $n^{th}$ element in $f$ contains the number of positron decays at the spatial location defined by the $n^{th}$ voxel. When the object is measured or imaged by the PET scanner, it detects the number of coincidence events (assume true coincidences only for now) along each of the $M$ LORs. Let these detected events be stored in a vector $g$.

This relationship between radioactive decay in $f$ and coincidence events in $g$ is represented by a matrix $A$, more commonly known as the system matrix. $A$ is an $M$ by $N$ matrix where $A_{nm}$ represents the probability that a positron emitted via decay at location voxel $n$ generates a coincidence event that is recorded along LOR $m$. However, based on knowledge of the physics behind the measurement system, system matrix $A$ can only be estimated in the form of $\hat{A}$ as assumptions are made about physical effects like detector geometry, positron range, detector sensitivity, depth of interaction, photon absorption and photon non-colinearity. While some effects like positron range and photon absorption are dependent on the subject being scanned, others are dependent on the
physical characteristics of the PET scanner such that $A$ is often decomposed into smaller independent components.

The reconstruction problem then sums down to solving equation 2.3 below:

$$g = Af$$

Equation 2.3

Where given $g$ and estimate of $A$ in the form of $\hat{A}$, $f$ can be estimated.

Mathematically, solving equation 2.3 to find $f$ given $g$ and $A$ presents itself as an inverse problem where simply matrix inversion would have sufficed to recover $f$. However, both $f$ and $g$ are very large ($10^7$ elements) such that inverting $A$ becomes an ill-posed problem, where a small error in $g$ would lead to a large error in $f$. As importantly, the process is computationally impractical as it requires for the computer to store and process a very large matrix. Instead of standard inversion techniques, image reconstruction algorithms that have been developed over the past 2 decades are used for the best estimate of $f$.

During or prior to image reconstruction, the effects describe in Figure 2.2 must be accounted for. Since equation 2.3 has a system matrix that accounts for factors that affect true coincidence events only like detector sensitivity or photon attenuation, other factors like scatter coincidences and random coincidences need to be included. Equation 2.3 thus evolves to equation 2.4.

$$g = Af + s + r$$

Equation 2.4

Where $s$ and $r$ vectors containing $M$ elements that represent scatter and random coincidences respectively.

The next section expands on the data corrections applied to PET data in a little bit more details as a couple of them are key to the motion correction algorithm that is presented further into this thesis.

2.4.1 Attenuation

When annihilation photon pairs interact with body matter from the subject, one or both photons are deflected or undergo absorption resulting in scattered coincidences or
attenuated true coincidences respectively. In PET, the attenuation of an LOR depends on the distribution, density and type of matter along that LOR. Whether the annihilations occur inside the body through the injected tracer, or outside through an external positron emission source makes no difference and the attenuation factor remains the same. This is because for a coincidence event to be recorded, the annihilation photons must travel through the whole of the LOR as the paths of each of them are combined. Hence, for every LOR, the attenuation of a true coincidence resulting from positron-electron annihilation is equal to the attenuation of a beam of γ-rays of 511 keV from an external source incident on the object along the same LOR. Attenuation correction in PET image reconstruction therefore implies estimating the probability of a photon to traverse the whole LOR without interacting with matter along that LOR.

A *transmission source* is typically used for calculation of attenuation correction, where an external source (a single photon emitter or positron emitter) is used to obtain 2 scans. The first one is called a *blank scan* when no object is present in the scanner. The second scan, called a *transmission scan*, is then made with the subject present in the scanner. The difference in count rates between the blank and the transmission scan is due the attenuated photons within the subject. The attenuation factors can then be reconstructed to calculate a *µ-map*: an image of linear attenuation coefficients. The µ-map can then be segmented into tissue classes (bone, soft tissue, air) in order to reduce errors in the attenuation correction factors by reducing noise in the image as well as reducing errors caused by scatter and emission contamination (Nuyts et al. 1999). The segmented image is then forward projected to give a better estimate of the attenuation correction factors.

In the absence of a transmission scan, attenuation correction factors may be derived by estimating the position of the subject from the emission data and then assuming a typical distribution of matter in the brain’s anatomy. This method is however known to be error prone (Hooper et al. 1996). In PET-CT equipments, the CT scan is used to derive the transmission scan. For this to happen though, the energy of the annihilation photons need to be scaled to 511 keV using a bilinear scale factor (Meikle et al. 2005). In the next generation scanners, PET-MR scanners, attempt is being made at deriving the
transmission measurements via the MRI scan. However, several difficulties linked to finding linear attenuation coefficients of energetic photons using data from images that do not have the same properties still proves to be a challenge (Kops and Herzog 2008).

### 2.4.2 Scatter

Scatter correction is regarded as the most complex part of PET data correction techniques (Bailey et al. 2004) and is now an even bigger concern due to the large increase in scattered events as data acquisition has moved from 2D to 3D. Techniques to estimate the amount of scatter within a subject include energy window manipulations where the scattered events generally have lower energy than true coincidences or by measuring the scatter from line or point sources within a scatter medium to eventually adjust a mathematical model, that represents the physics and geometry of the object and the scanner, that is then included into the reconstruction algorithm.

The most popular scatter correction algorithm used by the most common commercially available PET scanners are simulation based like the one by (Watson et al. 1996) where a simulation of the scattering process is carried out.

### 2.4.3 Randoms

Randoms are two detected single events within the same time window. For the majority of the events, both photons are not detected with the vast majority of detections being single events. It is due to low scanner sensitivity, typically between 0.3 – 3% efficiency for coincidences (Bailey et al. 1991; de Jong et al. 2007) and is further reduced by photon attenuation. Since single events are emitted in high numbers (as many as 10⁷ events per second on the HRRT), the probability of detecting two single events as an annihilation photon pair is relatively high. Random coincidences are typically measured by using a time-delay that is much greater than the coincidence window. A second dataset is thus generated that is devoid of any true coincidences and thus can contain only randoms. The delayed coincidences can then be used to estimate the randoms from the prompt or non-delayed coincidences. This method is however subject to statistical noise that leads to an increase in data variance. Measured delayed coincidences along with noise suppressing algorithms have been proposed to address the problem (Byars et al. 2005; Watson 2009).
2.4.4. Normalisation

Detectors vary in performance as their manufacture is never under the same conditions such that the relative sensitivity between detectors is not the same. Normalisation is the process via which the difference in sensitivity is considered in order to avoid artefacts in the PET image and inaccurate activity measurements as the relative sensitivity of each sinogram bin is not the same. A normalisation correction sinogram is therefore acquired that contains the inverse of the relative sensitivity of each bin. The coincidence sinogram is then multiplied by the normalisation sinogram to correct for the difference in sensitivity between bins. Methods for normalisation may be direct or component-based. They are however beyond the scope of this thesis and the reader is recommended towards these references for further information (Hoffman et al. 1989; Defrise et al. 1991; Bailey et al. 2004; Rodriguez et al. 2007).

2.4.5 Calibration

Calibration is ultimately used on PET images to generate absolute measurements of activity concentrations. It includes correction for overall detection efficiency of the scanner, correction for the branching ratio of the nuclide and allowing for the finite sensitivity of the scanner. The calibration scale factor effectively converts the image into units of absolute activity concentrations in kBq/ml. It is usually derived from a standard test object (usually a cylindrical phantom filled with water and a locally measured dose of $^{18}$F) whose image is reconstructed and the calibration factor tuned so that activity measurements from the well-counter or ion-chamber matches that obtained from the PET scanner. Changes in sensitivity are measured in daily quality controls via a uniform $^{68}$Ge phantom as according to (de Jong et al. 2009).

2.5 Image reconstruction techniques

PET image reconstruction techniques can be separated mainly into analytic versus iterative techniques. Filtered Back Projection (FBP) was the main method of image reconstruction since the very inception of PET as an imaging modality and, being computationally fast, dominated the image reconstruction scene until the 1990s, when
computational power became cheaper and more accessible. Then, iterative methods for image reconstruction that require more computational power compared to FBP, started to gain popularity to the point that they are now regarded as the standard method for reconstructing clinical data. The main reason as to why those iterative methods are preferred is because it allows incorporation of a statistical model that enables the nature of Poisson statistics linked to radioactive decay to be taken into account during reconstruction. This results in lower variance images enabling better visualisation of the anatomy and pathology.

2.5.1 Analytic reconstruction methods

In analytic reconstruction techniques, $\hat{A}$ needs to be simplified to enable the image process to be modelled as a set of line integrals. If the acquired and corrected data is stored in the form of sinograms (thus the Radon Transform), the inverse of the Radon Transform is most commonly achieved via Filtered Back-Projection (FBP) where the line integrals are projected back to their planes at their respective angles. Since this would result in a heavily blurred image. This is in turn because projection of the data, followed by a back-projection step, acts like a smoothing operation where-by the high frequency components of the true activity distribution are suppressed and the low-frequency terms over-emphasised. The use of a high-pass filter, typically a ramp or a Hamming filter helps in correcting for the incorrect weighting of the frequency components and suppresses the blurring effects on the true image. However, the recovered image is dominated by noise.

Prior to analytic reconstruction, the projection data requires arc-correction as well as correction for attenuation, scatter, randoms, normalisation and dead-time. Analytic methods are also devoid of a statistical model that leads to a lower Signal-to-Noise Ratio (SNR) as compared to iterative reconstruction techniques. The use of FBP also prevents the inclusion of known constraints like non-negative activity to the data. Because of amplified noise and mis-match between $A$ and $\hat{A}$, the images are generally prone to artefacts like ‘streak artefacts’ and are of poorer resolution than their iterative methods counterparts.
A review of the mathematics of FBP is found in (Defrise and Kinahan 1998). The characteristics of FBP that makes it desirable for use in image reconstruction are described in (Deans 1983; Natterer 1986; Barrett and Swindell 1996).

2.5.2 Iterative reconstruction techniques

Iterative reconstruction techniques do not need to assume the line-integral model as does FBP. The solution to equation 2.3 is instead obtained by obtaining an estimate of \( A \), then from an initial estimate of \( f \), iterating using a statistical model to estimate the distribution of the acquired projection data. Each iteration aims at comparing the measured data \( g \) with an estimate of \( g \) that has been calculated from an estimate of \( f \) and to minimise the difference between the measured \( g \) and the calculated \( g \).

(Shepp and Vardi 1982) described the Maximum-Likelihood-Expectation Maximisation (MLEM) algorithm which was the first iterative algorithm for reconstruction presented in emission tomography.

2.5.2.1 The MLEM algorithm

Let the measured coincidence data \( g \) contain \( M \) independent Poisson variables. The aim of MLEM is to find the maximum conditional probability or likelihood of the tracer distribution in image \( f \) that is likely to yield the measured coincidence data \( g \). This maximum likelihood is mathematically given by the product of all the Poisson probabilities (assuming independent events) of finding the true tracer redistribution that produces every measured data over the whole image and hence

\[
p(g|f) = \prod_{m=1}^{M} p(g_m|f) = \prod_{m=1}^{M} \left( \frac{\tilde{g}_m}{g_m!} \right) e^{-\tilde{g}_m}
\]

Equation 2.5

The Expectation Maximisation step then maximises this likelihood and is defined as:

\[
\hat{f}_{ML} = \arg\max_{f} p(g|f)
\]

Equation 2.6
From section 2.4,

$A_{mn}$ is the probability of detecting a true coincidence in the $m^{th}$ bin given that it has been emitted from the $n^{th}$ voxel. Hence, the total number of counts to be expected along a given LOR is the sum of the expected counts for that LOR from the total number of voxels within the image together with the expected number of random and scattered coincidences and hence,

$$E(g_m) = E(A_{mn}f_n) + E(r_m) + E(s_m)$$  \hspace{1cm} \text{Equation 2.7}$$

$$E(g_m) = \left[ \sum_{n=1}^{N} A_{mn} f_n \right] + E(r_m) + E(s_m)$$  \hspace{1cm} \text{Equation 2.8}$$

Where

$E(r_m)$ and $E(s_m)$ are obtained from the random and scatter corrections.

The MLEM algorithm, where an estimate of $f$ iteratively maximises on $p(g|f)$ is ultimately given by:

$$f_n^{(k+1)} = \frac{f_n^{(k)}}{q_n} \sum_{m=1}^{M} \hat{A}_{mn} \left( \frac{g_m}{\sum_{j=1}^{N} \hat{A}_{mj} f_j^{(k)}} + \hat{f}_m + \hat{s}_m \right)$$  \hspace{1cm} \text{Equation 2.6}$$

Where

$k$ is the iteration number,

$q_n$ is known as the sensitivity image and is defined as

$$q_n = \sum_{m=1}^{M} \hat{A}_{mn}$$
The algorithm is typically initialised with a uniform image at $f_n^{(0)}$

Provided that the estimates of $A_{mn}$, $r_m$ and $s_m$ are accurate, each iteration of MLEM produces an estimated image $\hat{f}_n^{(k+1)}$ that has a higher probability (and therefore a higher likelihood) than the previous iteration $\hat{f}_n^{(k)}$ of being the true distribution of activity. MLEM converges monotonically towards $f_{ML}$ but is terminated prior to convergence in practice. As opposed to FBP, MLEM maintains the non-negativity constraint as the factor by which $f_n^{(k)}$ is updated is always positive.

Several other flavours of MLEM have since been reported in literature. The one that is used for standard clinical data acquisitions on the PET scanner at the WMIC is 3D-OP-OSEM (Comtat et al. 2004). The LOR data is partitioned into $S$ ordered subsets, each with an equal number of projections, and the MLEM algorithm is then applied to the subset data only instead of the whole data. If the number of subsets is not too high, the image quality is preserved and the speed-up factor is $\approx S$. For example, if an EM-ML algorithm took up forward projection under 288 angles. For OSEM, that data could be split up into 16 subsets of 18 angles such that each iteration step would only require forward projections under 18 different angles. The speed-up factor would thus be 16.

2.6 The High Resolution Research Tomography (HRRT)

The ECAT High Resolution Research Tomograph (HRRT) is a dedicated human brain PET scanner developed by CTI Systems (now Siemens Medical Solutions) in Knoxville, Tennessee, USA in collaboration with the Max-Planck Institute for Neurological Research in Köln, Germany. The prototype model was installed in 1999 before 18 more systems were supplied to PET research centres across the globe between 2002 and 2008. The HRRT is not produced anymore and receives limited support from its manufacturer for improvement on the software front.

True to its name, the HRRT was essentially designed for high resolution research in brain imaging in a bid to develop PET scanner technology. The HRRT has a few innovative features which enabled its developers to get valuable insights on the benefits and difficulties that upcoming PET technologies would offer in the near future. It also allowed
further improvements in image reconstruction methods based on its unique characteristics which are briefly outlined below.

2.6.1 Hardware

The HRRT consists of 8 flat panel detectors arranged in an octagonal structure as shown in Figure 2.4 below. Each detector head is 252 mm long by 195 mm wide and contains 117 detector blocks arranged in a 13 x 9 array. Each block is a cuboid with a square face of 19mm that consists of a first layer of LSO crystals followed by a second layer of LYSO crystals that are both 10mm thick. A 5.5mm thick light-guide then follows. Each layer of crystals contains 64 individual elements arranged in an 8x8 grid as shown in Figure 2.5a, with each element being a square-faced cuboid of width 2.1mm and depth of 10mm. The 9x13 array of blocks is coupled to a 10x14 array of photomultiplier tubes (PMTs) in a quadrant sharing design, meaning that each PMT is optically coupled to up to 4 blocks as shown in Figure 2.5b

![Image](image.jpg)

**Figure 2.4:** Picture of the HRRT with the casing open showing the detector heads on the left hand side and corresponding schematic on the right hand side.

It has a transaxial FOV of 312 mm and its axial FOV of 252mm is longer than that of most PET scanners, thus allowing the entire brain to be imaged from a single bed position. With
a limited entry port of 350mm, whole body scanning is unachievable. For acquisition of a transmission scan, it uses a $^{137}$Cs point source of 1.1 GBq that is encapsulated in a Tungsten shielding that also forms a collimator, producing a narrow fan beam of γ-rays when the collimator is open. It is kept behind the gantry, outside of the FOV when not in use and during a transmission scan, moves to in between the detector heads and the edge of the patient port and rotates through 270° around the subject, with steps along the axial direction. For a more detailed review of the HRRT’s hardware, the reader is directed towards the detailed information from (Knöß 2004).

The high spatial resolution of the HRRT is achieved via:

(i) The high number of small finely cut crystal detection elements (2.1mm square-face as opposed to 4mm square-face in clinical systems) that is achieved via the use of a flat panel design.

(ii) The quadrant-sharing block design that enables the use of a reduced number of PMTs over the finely cut crystal elements. (2.1mm square-face compared to 4mm square-face in clinical systems)

(iii) The presence of 2 layers of crystals that allows for Depth of Information (DOI) to be recorded. DOI helps in reducing parallax errors that occur due to the uncertainty at which the annihilation photons interact in the crystals as shown in figure 2.5c and helps with the loss of deterioration from the centre of the FOV, moving out radially.

(iv) The LSO/LYSO scintillators that have a higher light output and a faster decay time than scintillators like BGO (Daghighiam et al. 1992).
**Figure 2.5:** Detector block design for the HRRT. (a) 8x8 grid of LSO/lyso crystal layers connected to light guide and PMTs in quadrant sharing design. (b) Detector blocks of size 13x9 for one detector head and associated PMTs. (c) The allowance for Depth of Interaction (DOI) in a crystal of the HRRT that minimises parallax error in LOR measurement.

### 2.6.2 Software

The standard manufacturer software supplied for the HRRT allows for reconstruction of dynamic images with all the corrections mentioned in Section 2.5. The standard algorithm used in clinical research at the WMIC is the 3D-OSEM algorithm by (Comtat et al. 2004). Attenuation correction is derived from the transmission scan through the reconstruction of a segmented $\mu$-map that includes: scaling of attenuation coefficients from 662 keV to 511 keV, acquisition of a blank scan for compensation of detector sensitivities and compensation for activity in the FOV by acquiring a shifted mock transmission data (Sibomana et al. 2004). Scatter correction is achieved via the Single Scatter Simulation method of (Watson 1999). Randoms are estimated from delayed coincidences as per (Byars et al. 2005; Van Velden et al. 2008a). The direct method for normalisation is performed using a rotating rod source and correction for dead-time is

54
also done prior to calibration in kBq/ml using data from measurements taken off a cylindrical phantom filled with water containing $^{18}$F. Figure 2.6 shows a flowchart of the key processes applied to the listmode data during the reconstruction process in order to obtain an emission image. The flowchart includes only those processes that are used later on in chapter 4 for developing the motion correction algorithm.

Due to the inherent research nature of the HRRT, scanner software and acquisition systems are open to the user. In fact, in part due to limited support from the manufacturer, the 19 users of the HRRT have grouped themselves together to form the HRRT Users’ Community, a very organised and active group that is acknowledged and supported by Siemens Medical Solutions, where the members combine their knowledge and expertise to develop and support the HRRT software.

The processes described above for image reconstruction is done via scripts, some of which are summarised in flowchart Figure 2.6. The fact that these scripts are available for modification and work independently is one of the main reasons why the motion correction algorithm presented in chapter 4 was successfully developed. As a preamble to what is dissected in more details in chapters 3 and 4, one of the reasons why patient movement causes image artefacts is due to emission-transmission mismatch, leading to incorrect attenuation and scatter correction factors. In-house developed scripts like split_jcm_1.3 allows researchers to use the temporal flexibility of listmode data to define frames within which there is minimal motion within the frames. The attenuation correction sinogram is obtained by the reconstruction of a $\mu$-map from the transmission image. The HRRT reconstruction software has been structured in such a way that several $\mu$-maps may be reconstructed such that the appropriate attenuation correction sinograms will be available for the respective emission data in a dynamic reconstruction.
Figure 2.6: Flowchart of key processes that are run via scripts on the HRRT reconstruction software. In **blue** are those outputs that are modified/used to alter the current reconstruction process to include the developed motion correction algorithm. In **green** are the processes that are used in the modification process via their respective scripts that are in **red**.

The HRRT Users’ Community also gears its efforts towards improving reconstruction software available for the HRRT like the speeding up of the 3D-OP-OSEM reconstruction algorithm (Hong et al. 2007), or the incorporation of a resolution model for the higher resolution and reduced variance in reconstructed images. Recently, the HRRT Users’ Community has shared a 3D FBP reconstruction algorithm (Van Velden et al. 2008a; Van Velden et al. 2008b) as well as an improved attenuation correction method (Sibomana et al. 2009).

For a more detailed review regarding the construction and characteristics of the HRRT, the reader is referred to the references of (Wienhard et al. 2002; Knöß 2004) which are specific to the prototype version of the HRRT. The compared final design is in (de Jong et al. ...
al. 2004). The performance characteristics are reported in (de Jong et al. 2007) while comparison between HRRTs can be found in (Sossi et al. 2005).

2.7 Summary

The purpose of this chapter was to understand the basic concept that allows one to obtain the measurement of the spatial distribution of activity within the brain (or any part of the body/object) using PET physics. The principles of image reconstruction in emission tomography were also reviewed as well as the importance of correction processes like attenuation and scatter in image reconstruction. The HRRT is touched upon, both from a hardware and software point of view.

The important notes from this chapter are;

(i) Acquisition of PET data in listmode format is a desired feature as the timing information makes it an added benefit for application of efficient motion correction methods (looked upon in more details in Chapter 3)

(ii) The attenuation correction sinogram can be obtained by reconstruction of a μ-map that is in-turn derived from the transmission image.

(iii) The scatter correction sinograms are derived using the attenuation correction sinogram.

(iv) The HRRT is a dedicated high resolution brain PET scanner that was design principally for clinical research purposes (and as a prototype model for upcoming new technologies PET). It is open-sourced such that researchers have access to the lowest-level of coding of the scanner’s acquisition and reconstruction software, allowing improvements on that front.

(v) In the HRRT reconstruction software, the process for obtaining the attenuation and scatter correction sinograms can be scripted and run as separate entities. This, coupled with the fact that the software is open-source allows the inputs and outputs of the dataflow shown in figure 2.6 to be modified in order to incorporate algorithms like motion correction or resolution modelling.
The author set this chapter with a perspective view of how the motion correction algorithm is implemented. The materials provided in this chapter may seem very raw at present reading but is valuable information for any reader who wants to probe into the low-level details as to why the motion correction algorithm was ultimately devised in this manner.

The next chapter looks into why motion correction is a problem in PET when imaging at a high resolution with a scanner like that of the HRRT and the different methods developed in literature to address the problem. Based on literature, practical problems observed in daily clinical research at WMIC, and analysis of the benefits available to researchers in terms of ability to improve on data processing, an improved method for motion correction is ultimately proposed.
Chapter 3
Review of Motion Correction Methods

3.1 Introduction

A typical \(^{18}\text{F}\)-FDG PET scan in clinical research lasts for about 70 minutes. \(^{11}\text{C}\)-MDL or \(^{11}\text{C}\)-DASB scans have a 2-hour duration. It is unreasonable to expect patients to remain still during these long scan times, especially in dementia studies where besides suffering from movement disorders, the level of discomfort in older patients is higher and their tolerance to it lower. Head restraint, though simple, is more of an option towards minimisation of motion rather than eliminating motion. Use of extreme fixation methods like a thermoplastic mask with a bite piece would be unethical. Moderate forms of restraint that have been tried are moulded plastic masks, vacuum-lock bags, orthopaedic collars and straps. They have been found not to eliminate movements completely and, depending on the type of masks used, translations in the range of 5 mm – 20 mm, occurring typically along the transaxial axis of the PET camera (and de facto, the subject), and rotations between 1° - 4°, occurring typically around the axial axis, have been reported (Green et al. 1994; Lopresti et al. 1999).

A classification of the different types of movements observed during a scan has been proposed by (Lopresti et al. 1999). However, classification of movements into random and non-random types is rather confusing in the author’s point of view. For example, they define coughing as being a non-random type of movement. In statistics, an event is regarded as random depending on its stochastic nature. Hence, an event like coughing is more stochastic than breathing for example, which makes the classification of movements rather confusing. It is probably easier to classify the types of movements into:

1. Frequent (occurring many times within a frame), irregular movements that can occur in any direction, usually in high numbers and small amplitudes. Examples are breathing, fidgeting or tremors. Statistically, they are known to form a
distribution around a mean value for each Degree of Freedom (DoF) and contribute to a spatially homogeneous loss of effective image resolution.

2. Infrequent movements that may or may not occur at all within a frame. They occur at a lower frequency as compared to frequent movements, but are larger in amplitude and most often, have specific directions. Examples of such movements include coughing, sneezing, attempts to relax body muscles (leg-crossing, lower-torso repositioning), response to pain and sudden awakening. Such movements have both, spatial and temporal impacts on the image as they tend to cause gross misalignment between transmission scans and emission scans. They, de facto, cause changes in attenuation and scatter correction factors that eventually lead to incorrect sinogram data and shifting of the location of functional activity within the region.

3. Drift – defined as slow gradual motion as a patient tries to slowly relax his/her posture, falling off to sleep, unknowingly adjusting his/her head to a more suitable position or recoiling in his/her sleep due to feeling cold. Drifts can also be encountered post a large discrete motion, where the subject makes an attempt to return to his/her initial position with a slow gradual movement that is often repetitive. Drifts can take several minutes to complete where the final head position is different to its initial position.

These movements are known to degrade the effective spatial resolution of the PET scanner. (Green et al. 1994) have measured and analysed head movements in normal controls with and without head restraint. To have a general idea of the impact of head movements on the resolution of the scanner, the authors have expressed its effective resolution as a combination of the FWHM of intrinsic resolution of the scanner in quadrature with the FWHM of the distribution of the patients’ head positions around a mean value over the scanning interval. This is expressed mathematically by equation 3.1 below.
\[ \Delta_{eff} = \sqrt{\Delta_{tomograph}^2 + \Delta_{motion}^2} \]  
\text{Equation 3.1}

Where, 
\( \Delta_{eff} \) = The effective resolution of the PET scanner;
\( \Delta_{tomograph} \) = The FWHM of the intrinsic resolution of the PET scanner;
\( \Delta_{motion} \) = The FWHM of the distribution of the patient’s movements around a mean value.

The authors however point out that the use of the term FWHM implies that they have scaled the standard deviation of their measured movements so that it approximates to a Gaussian distribution. It neither means that the distribution of the patients’ movements was Gaussian, nor that the quadrature combination between the latter and the scanner’s resolution yields another Gaussian distribution. It only highlights the fact that as tomographs evolve to a higher resolution, the component due to motion becomes more pronounced. The scientific impact of this spatial degradation is that it hampers the ability of these high performance scanners to measure small changes in radioligand binding coming from small brain structures like the nuclei as well as to better delineate the patterns of the cortical gyri. In order to observe the benefits of PET imaging at high resolution, the motion problem needs to be addressed.

### 3.2 History

Motion correction strategies made their first appearance in the early 1990s. They were effectively automatic, rigid-body co-registration methods (Woods 1992; Friston et al. 1995), applied post-reconstruction, to correct for inter-scan motion: that is, motion between scans of the same patient, acquired under the same scanner at different points in time. In (Woods 1992), the basic idea of the algorithm is that if a registration between 2 PET slices is successful, the relationship between the voxels of the source image and that of the target image is a simple multiplication factor. The latter should be uniform across voxels of both images, meaning that the coefficient of variance across the image should be constant; else the volumes have been misaligned. In their research, 2 series of
PET slices were acquired axially and were stacked to obtain the sagittal view using simple linear interpolation. Masks of the ratio volumes (of the resliced volume to the reference volume) were then obtained and checked for homogeneity. If the ratio was found to be varying, the resliced volume would be iteratively aligned using tri-linear interpolation and known translation and rotation parameters until the final ratio volume is homogeneous. The main aim of this study was to use this registration algorithm to alter the settings of the bed and gantry of the scanner such that a new scan of the same patient would be acquired in the same space as the former scan.

Motion correction is however different from aligning two different scans. It deals with motion that has occurred during the scan. In clinical research, as a strategy for motion correction, such a method is incorrect as they make the assumption that the activity within brain regions is constant and that the emission scan is aligned to the transmission scan during acquisition. Motion correction strategies in emission tomography have been devised for the very reason that there is an emission-transmission mismatch as the patient moves during the scan. In fact, (Woods 1992) does concede in the evaluation of his method that it would not solve for motion that occur during a scan and have proposed to acquire several attenuation scans in different positions and to interpolate between the transmission images to derive the appropriate one.

3.3 Motion Detection Methods

The main conclusion from post-reconstruction frame-by-frame realignment is that it has limitations regarding its use in a clinical setting, as it does not solve the problem of a mismatch between emission and transmission image, causing the use of wrong attenuation and scatter correction factors in image reconstruction. It also does not correct for motion that has occurred within a frame. For efficient motion correction, a good method for motion detection and quantification is crucial. Quantification of motion can be done via independent external equipment which explains the era of 1995 till present, in which research groups have been focusing on the design and implementation of instrumentation systems that would track and measure movements during PET studies. Systems that used electromagnetism had to be avoided because of the Eddy currents
generated in the PET gantry system, acoustic devices were likely to interfere in neurological studies (Goldstein et al. 1997) as they would activate sites like the superior geniculate of the thalamus, Brodmann areas 41 and 42, and areas of the superior temporal gyrus. In that respect, 3 main motion detection devices have been found in literature. It started with a camera based surveillance system by (Picard and Thompson 1995) that culminated to the POLARIS system by (Lopresti et al. 1999) via the optoelectronic system by Goldstein et al (1997). In the camera based system of Picard and Thompson, 2 Charged-Couple Device (CCD) are orthogonally mounted onto the gantry of the PET scanner and 3 Light-Emitting Diodes (LED) are placed on the source so that they can be tracked by both cameras and provide the 6 Degrees of Freedom required to transform the patient’s position with movement. The images of the 6 LEDs, that appear as 6 matrices that are 6x6 in size are stored in frame buffers that are recorded at a frequency of 30Hz. The centroids of the 6 LEDs are then monitored and compared to the centroids obtained in the previous frame. If they are different, the position of the new centroid is updated on the matrix. The method was designed mainly for patient repositioning so that they were imaged in the same space as their previous scans. It was found that the system could show motion depending on the subject’s physiognomy. (Goldstein et al. 1997) used 3 micro-lamps fixed to the patients’ heads and their positions determined via triangulation of the information output by the electro-optical position sensitive detectors. The objectives of this paper was to accurately monitor transformations due to movements, be as non-invasive as possible and not interfere with the operation of the scanner or scanning operations. Results show good response of the system in terms of accuracy and linearity of measurements of translations and rotations and the authors believed that implementing the system on a motion correction algorithm would bring positive results. The system is however prone to measurement errors if the micro-lamps are not securely fixed to the subject. The most acknowledged motion detection device is the POLARIS by (Lopresti et al. 1999). It consists of an infrared (IR) optoelectronic motion tracking device that tracks the position and orientation of four IR reflective spheres (four used in the works of Lopresti et al. but more can be used) that form up a tool interface of precise geometry. The latter however has to be in a working volume defined by the position sensor and the IR reflective spheres need to have their
relative positions to each other fixed. The position sensor uses triangulation and pattern recognition techniques for the tracking. However, while the system is good at tracking the tool, the latter needs to be fixed to the subject in order to measure cranial motion. In practice, this was found to be difficult. At the WMIC for example, the tool design is fixed to a neoprene cap, and it was observed that the skull would move while the cap (and therefore the tool) would remain still. At one of the HRRT’s Users Community Meeting, one research centre reported that their tool interface was sensitive to skin movements that did not necessarily imply skull movements. Several ways of attaching the tool to the patients head have been tried, with the most common ones in literature being wet caps, helmets, goggles and band-aids. Uncertainties over how efficient these fixations are over long scan durations in patients’ studies still prevail (Herzog et al. 2005). Use of more ‘extreme’, but arguably more efficient, methods like skin glue has been reported in animal studies only (Jin et al. 2010). Potentially new methods of motion detection are currently under exploration in the field of Structured Light (SL) Tracking (Olesen et al. 2010). The rationale for using such a system is that it does not require external markers to be fixed onto the patient as it uses LEDs in the Near Infrared (NIR) region to create and track a 3D reconstructed surface of the object. The idea of SL for motion detection is promising, but still very work-in-progress as part of the light being thrown onto the patient is lost through absorption and scatter, which leads to a loss of sensitivity of the system.

3.4 Motion Correction Algorithms

Based on the motion detection system available, different motion correction algorithms have been proposed in literature. These can be divided into 2 main groups: (i) Image-realignment approaches: where the correction for the measured motion is applied to images and (ii) data realignment approaches: where the motion correction method is applied to the LORs prior to reconstruction of the emission data.
3.4.1 Image-realignment methods:

As mentioned earlier in the introduction section, the very first image-driven methods (Woods et al. 1992) were devised with the aim of realigning scans for the same patients acquired at different points in time. These scans were mainly static images where the accumulation of a tracer like $[^{15}\text{O}]-\text{H}_2\text{O}$ over a short period of 70s was used in blood flow studies. Application of this same principle to dynamic studies where motion is to be corrected for within a scan is not viable as the method does not account for mismatch between the emission and transmission data. Only one research group has implemented and published the method, probably under the constraint of being the only option towards improving their data (Mawlawi et al. 2001). In that study, the frames were first denoised using wavelet filters. One denoised frame was then selected as reference, to which the other frames were co-registered and the transformations mapping the movement from those frames to the reference frame were saved (TM$_i$, where $i$ refers to the denoised scan being co-registered). The reference frame was then mapped onto the corresponding MR image and this transformation too was saved as TM$_2$. The original scans were then mapped onto the MR by a straight combination of TM$_i$ and TM$_2$. The wavelet filtering was deemed an important step to make the co-registration step more robust as short early frames are inherently very noisy and thus, a source for co-registration mistakes. Moreover, registration errors can also find source from tracer redistribution at the early vascular stage following injection of the radiotracer (and for other tracers, at later stages during a scan as well). While the denoising process aims at reducing variance in the data, it is not relevant to redistribution effects happening within the brain. However, it may improve on the goodness of fit of the data in the co-registration process as the noise in the system has been reduced. Despite its late publication (other research groups have used improved methods and published before this one), outlining the authors’ work is still important as the method has always been used as the basis for what can be defined as ‘conventional Frame-By-Frame (FBF) realignment’, against which improved image-based methods have been compared, albeit without application of wavelet denoising filters.
One of the main drawbacks of realigning images post-reconstruction is that it does not account for the emission-transmission mismatch. To address this problem, (Andersson et al. 1995) proposed to reconstruct the emission data first without attenuation and scatter to define the movements during the scan using image registration. Assuming that the first emission frame \((EM_1)\) contained no movements with respect to the transmission image \((TR)\), the movements between \(EM_1\) and the subsequent frames \((EM_i)\) would then be quantified via co-registration and the transmission image mapped to \(EM_i\) using the inverse of co-registration transformation parameters calculated. Each \(EM_i\) is then reconstructed, this time with the appropriate attenuation and scatter correction factors and then resliced to the reference \(EM_1\) using the forward transformations. The authors applied this algorithm to head and cardiac movements in human patients and found that the method was able to detect and correct for movements of the order of at least 5mm.

The method for detecting motion in (Andersson et al. 1995) is effectively the reconstruction of emission frames without attenuation and scatter. In this method, the frames are defined \textit{a priori} such that the problem of motion within a frame is not addressed. One of the advantages of having an external movement detection system is that it can detect for motion at any point during a scan, hence addressing the problem of within frame motion. (Menke et al. 1996) used a video camera based system to measure movements. Based on those movements, a Wiener filter is used to deconvolve the images that are regarded as having been blurred by the movements measured during the scan. The shape of the deconvolution operator is chosen according to the movements observed. However, deconvolution is known to amplify noise in PET data which is inherently noisy. Furthermore, in the presence of significant noise, a spatially variant filter needs to be used that increases computational costs and the potential of introducing artefacts like Gibb’s artefacts. For these reasons, the method was not acknowledged by the PET community. A more sensible and acceptable alternative was proposed by (Picard and Thompson 1997) which came to be known as the Multiple Acquisition Frames (MAFs) approach. Using their video camera system, they monitored the movement of the subject and based on a preset threshold, the system would acquire the PET data in a new frame. Those frames would then be reconstructed and co-registered to a common space based on the movements recorded by the video camera.
system. There are 2 main flaws in this method. The first one is that the new frames are chosen according to a preset threshold. If this threshold is too high, the amount of motions within frames would increase. On the other hand, if too low and if the patient moves a lot, this would result in a high amount of low statistics frames. Short, low-statistics frames are very noisy which can ultimately compromise successful co-registration. Moreover, the use of an algorithm like OP-OSEM for reconstruction would induce a bias in the measurement of activity counts within regions of interest even if the error due to motion is catered for (Reilhac et al. 2008). If reconstructed with FBP, the number of noisy images with streak ‘artefacts’ would increase. The second flaw of this motion correction algorithm is that it does not correct for emission-transmission mismatch as it uses the original transmission image to derive the attenuation correction factors. The reason advanced by the authors for doing this is that a portion of the attenuating materials present within the FOV do not move with the head (example: head restraint). These would need to be segmented out before applying the transformations to the brain image only and such a system was not yet in place. (Fulton et al. 2002) on the other hand, used the MAF approach with the POLARIS system, allowing attenuating media within the FOV to move with the head when matching the transmission data to the emission image. They observed that doing this still produced better results than observed by Picard and Thompson. They discussed that one of the reasons for disparities between a motion-free scan and one corrected with the MAF approach is the incorrect transformation of the fixed attenuating media during transmission-emission image matching.

(Wardak et al. 2010) addressed this problem by manually segmenting out the head-holder from the transmission image. Their motion correction scheme from then on follows that of (Andersson et al. 1995) and yields transmission images that are each aligned to the emission images reconstructed without corrections for attenuation and scatter. The segmented head-holder is then added back to these transmission images before deriving the attenuation and scatter correction factors for each corresponding emission frame. The work by (Wardak et al. 2010) had a more clinical than methodological perspective as they evaluated the benefits of the method against uncorrected data in an [18F]-FDDNP study, where the presence of β-amyloid plaques and neurofibrillary tangles were
compared between 12 AD patients and 9 age-matched controls. A Logan analysis was carried out with the cerebellum used as reference region to estimate Regional Distribution Volume (DVR) in different regions of the brain. Motion correction allowed the observation of significant differences in DVR between AD and controls (p<0.05) in the frontal, parietal, posterior cingulate, Medial Temporal Lobe, Lateral Temporal Lobe and global regions when compared to the same data without motion correction. They also observed a decrease in variability of the DVR values by 18% and increase in separation between the mean DVRs between the 2 groups.

3.4.2 Methods applied to LOR of raw data:

With more sophisticated instrumentation systems for motion detection, motion correction algorithms were extended from image-based to event-based, where motion correction is applied to the LORs prior to reconstruction. The recorded movements are effectively used to backproject the LORs to the original position where they should have been detected had the movements not occurred. This is illustrated in Figure 3.1

![Figure 3.1](image)

**Figure 3.1**: Illustration of movement of an event within the brain. The event (yellow star) that should have been detected along LOR $j$ is effectively detected along LOR $i$. The movement is measured externally using a motion tracking device to restore the event to its original position.
(Menke et al. 1996) event-based approach consisted in rebinning the LORs according to the motion detected from 2 Infra-red sensitive CCD video cameras that tracked 3 circular pieces of infra-red reflective tape fixed to the subject to provide the 6 Degrees of Freedom motion parameters for defining the subject’s head orientation and positioning. Knowing the geometry of the scanner, they derived large Look-Up Tables (LUTs) that allowed them to infer each movement measured to a certain repositioning of the LORs. Due to hardware limitations, they were unable to correct for normalisation factors. Not accounting for normalisation factors results in ring artefacts.(Buhler et al. 2004) addressed the normalisation correction issue and observed reduced artefacts. They also observed another issue linked with out-of FOV LORs that is discussed further below. They believe that algorithms designed to assess the full 3D nature of the LOR need to be developed to fully correct the artefacts. (Jones 2001) achieved on-the-fly physical corrections by creating a multi-level architecture for processing a detected pair of events. Level 1 caters for the correction factors like normalisation and dead time. The detector pair events are stored as scintillator indexes which are converted to 3D coordinates in Level 2. The transformation of the coordinates to address motion correction is done on Level 3. Level 4 rebins the data. In the final level, the data is histogrammed according to the physical correction factors calculated in Level 1.(Bloomfield et al. 2003) achieved motion correction by tracking the motion data with the POLARIS and aligning this data with the listmode data by generating unique 4-bit pseudo-random numbers that are inserted into the listmode data stream post acquisition. The LORs are then converted from sinogram-bin element to a 3D vector that is defined by the geometry of the scanner. These are then motion corrected according to the detected data and the LORs converted back to sinogram-bin elements, thus updating the sinograms. Normalisation correction factors and dead time correction are then applied and the emission data is finally reconstructed with the attenuation and scatter correction factors derived from the transmission image, which was the dedicated reference image. They assessed the performance of their motion detection and correction method by simulating for discrete motion using point and line sources and for continuous motion using the Hoffman Brain Phantom. They reconstructed the uncorrected and motion-corrected data and compared via profiles the closeness of the corrected data to the initial position of the source. For
the brain phantom, they assessed the resolution degradation between uncorrected and motion-corrected phantom with an image of the latter when scanned in a stationary position. They also applied the method to an in-vivo $[^{11}\text{C}]$-Raclopride study on normal volunteers, where the data was acquired as a dynamic series and compared Time-Activity Curves between corrected and uncorrected data. Besides the clear benefits of motion detection and correction in terms of resolution, (Bloomfield et al. 2003) also observed that event-based motion correction allowed them to correct for the most severe continuous motion that they had simulated for and still obtain satisfactory results, as long as the sampling rate of the motion detection instrumentation was more than the rate at which the LORs were being displaced (They set the POLARIS to work on 5 frames per second, while the LORs were being displaced at 5.9 per second in the continuous motion experiment; the POLARIS can be set to a working frequency of up to 20 Hz). They also briefly highlight a couple of limitations of their method which has been addressed by other research groups. These limitations are:

1. An event that is meant to be within the FOV will go undetected as it goes out after motion. This would result in incomplete sinograms and therefore artefacts in the reconstructed images.

2. Conversely, an event that has moved into the FOV originates from non-existing detector pairs and now contributes to the detected data.

These 2 limitations occur in the axial direction as shown in figure 3.2a. However, for the HRRT (and scanners that have gaps in between detector heads), they also occur in the transaxial direction as illustrated in figure 3.2b.
Figure 3.2: Illustration of LORs going in and out of the FOV as a result of motion in (a) the axial view and (b) the transaxial view. The red colours indicate the original position of the head and the LORs displayed. The blue colours indicate the position of the head after movement in the indicated direction. (1) An event that should have been detected across the solid red LOR goes undetected as the solid blue LOR is lost. (2) An event that should not have been detected as the dashed red LOR goes through the gap of the scanner or out of the rings is effectively detected as the dashed blue LOR is recorded by the scanner crystals.

(Thielemans et al. 2003) addressed the first problem by comparing the number of counts in the motion corrected sinogram bins to the number of counts in the reference sinogram bins. If the motion corrected one has less than the reference, a weighted scaling factor is used to correct for the missing counts. The scaling factors are weighted to avoid noise amplification when the scaling factors are too small.

Like (Thielemans et al. 2003), (Buhler et al. 2004) addressed the second limitation by simply discarding the LORs as they would not have been detected anyways. To compensate for counts that go undetected, they scale the normalised number of counts in each sinogram bin by the ratio of the frame time to the time during which the sinogram bin would have been detected within the FOV.

Focusing on the second limitation, while discarding the LORs that happen to fall out of the FOV does not lead to image artefacts, it has been shown by (Rahmim et al. 2004) that it still contains useful information that can be used to enhance the Signal-to-Noise Ratio.
(SNR) of the reconstructed images. In fact, the alternative method proposed by (Rahmim et al. 2004) addressed both limitations by using the EM algorithm with a modified system matrix for image reconstruction. In histogram-mode, neglecting normalisation, attenuation and scatter factors, the conventional EM algorithm may be written as Equation 3.2:

$$f^{k+1}(j) = \frac{f^k(j)}{\sum_{i=1}^{l} p_{ij} \sum_{b=1}^{l} p_{ij} f^k(j)} \sum_{i=1}^{l} \frac{p_{ij} n(i)}{\sum_{b=1}^{l} p_{ij} f^k(j)}$$

Equation 3.2

Where \(f^k(j)\) and \(f^{k+1}(j)\) are the previous and new estimates of activity distributions within the reconstructed image; \(n(i)\) is the number of events detected along bin \(i\); \(p_{ij}\) is the system matrix which denotes the probability that an emission from voxel \(j (j = 1,2,3, \ldots J)\) is detected along bin \(i (i = 1,2,3, \ldots I)\).

The modified equation translates to Equation 3.3:

$$f^{k+1}(j) = \frac{1}{T} \int_{t=0}^{T} \frac{f^k(j)}{\sum_{i=1}^{l} p_{ij} \delta_t^i} dt \sum_{i=1}^{l} \frac{p_{ij} n(i)}{\sum_{b=1}^{l} p_{ij} f^k(j)}$$

Equation 3.3

Where,

\(T\) is the total frame duration and

$$\delta_t^i = \begin{cases} 1 & \text{if the bin } i \text{ was detectable at time } t \\ 0 & \text{otherwise} \end{cases}$$

Normalisation correction can then be applied as a pre-correction factor or as a component of the system matrix element.

The approach is also presented for incorporation in listmode-based image reconstruction as opposed to histogram-based. Applying motion correction in listmode format implies that the data can be corrected as a continuous stream as opposed to when they are binned into sinograms and are not continuous anymore. Furthermore, the calculation of
the motion averaging term within the system matrix can be sped-up by calculating the latter in image space rather than projection space.

Recently, an event-based motion correction algorithm that is applicable to fully 4D image reconstruction algorithms was implemented where PET superset data is used to store motion compensated data measured from the POLARIS (Verhaeghe et al. 2010).

### 3.5 Comparison of motion correction algorithms in literature

There exist only a few papers in literature that explicitly compare the performance of different motion correction algorithms.

(Fulton et al. 2004) compared the MAF method to their earlier described LOR rebinning method by simulating movements on the Hoffman brain phantom. Since essentially continuous motion was manually applied to the phantom, 20 frames were acquired, each of 90 seconds duration and the transformations between the reference position and the position at the start of each frame. The MAF method used included correction for attenuation and scatter.

They observed that both, the MAF method and the LOR rebinning method, showed reduced errors when compared to a motion-free acquisition. The LOR rebinning method performed better than the MAF method (from profiles drawn across the phantom, the sum of squared difference between LOR rebinning and motion-free scan was 3.9 x 10^7 cps², as opposed to 5.2 x 10^7 cps² for the MAF method and 37.9 x 10^7 cps² for the uncorrected data), largely because the motions simulated were continuous in nature, which accentuated the amount of intra-frame motion.

(Montgomery et al. 2006) carried out a more comprehensive evaluation of the methods available for motion correction (Montgomery, Thielemans et al. 2006). In fact, it is the only quantitative comparison available in neuroimaging with PET, where the authors evaluate the performance of FBF realignment method as laid out by (Andersson et al. 1995) against methods that use external motion tracking for detection and correction. 8
volunteers were scanned with $^{11}$C-Raclopride and their Binding Potential (BP) were assessed over 2 steady-states of the scan. Test-retest indices were also used for quantitative evaluation of the efficiency of the motion correction methods. Motion tracking measurements were taken with the POLARIS. The methods under investigation were:

(i) FBF realignment as according to (Andersson et al. 1995); where the emission data was first reconstructed without attenuation and scatter, then denoised using wavelet filters and co-registered to a reference 25 minutes emission image using Mutual Information as similarity measure. The transformation parameters were then used to project the transmission image into the respective emission frames and the data reconstructed with the appropriate attenuation and scatter corrections.

(ii) Motion Tracking frame-by-frame (MTfbf) where the POLARIS is used to track movements of the head. The average position of the head is calculated for every frame and the transmission image realigned to that average position.

(iii) Motion-Tracking (MT) as per (Bloomfield et al. 2003), where the LORs recorded from the data acquired in listmode are moved to the position of the reference transmission scan via motion tracked using the POLARIS.

(iv) LOR Motion Corrected frame-by-frame (LMCfbf), as opposed to the above processes, is a 2-stage process. In the first step, the LORs within the frame are realigned to an average position within that frame via the POLARIS data, moving the attenuation map to that average position and reconstructing the images. In the second step, all the reconstructed frames are reverted back to the original position of the transmission scan reference frame. In this way, the magnitude of movements with respect to the transmission image is limited to the individual frame only. This makes the scatter estimation more robust. It has been shown by (Thielemans et al. 2006) that the scatter distribution changes with movements and that the corresponding bias in the reconstructed image increases with the magnitude of the movements.
In all cases, motion correction was observed to improve measurements by showing reduced test-retest variability as opposed to the uncorrected data. Results, drawn across 5 of the subjects in regions of the Dorsal Striatum (DS), Ventral Striatum (VS) and Cerebellum (CB), showed improved TACs, especially for MT and LMCfbf against the raw data and the frame-by-frame realignment methods. The improvement that LMCfbf brings to the data as compared to MT appears marginal on the TACs but is quite clear on the quality of the reconstructed image. In an ANOVA test on the effect of the method used for realignment with no interaction of the ROIs, a statistically significant reduction in the rate of change of activity in the VS was observed when comparing LMCfbf with the raw data (p=0.08) and when comparing LMCfbf with FBF (p=0.04). However, no significant changes were observed when comparing the other methods (MTfbf, MT) arguably showing little benefit in performing event-based motion correction as opposed to image-based realignment with consideration for attenuation and scatter. Noise levels across all 3 regions were also reduced. Test-retest indices showed significant reduced variability in the VS only, with no significant changes in the DS. More precisely, the Intra-class Correlation Coefficient (ICC) for the VS increased with all motion correction methods indicating better comparison of differences between subjects. However, there seemed no distinction between the methods used for motion correction. In the DS region, the ICC results looked even more random.

The authors underline the fact that MTfbf was inferior in performance to LMCfbf mainly because MTfbf does not correct for within-frame movements. In fact, they state that in the absence of such movements, MTfbf and LMCfbf images were identical.

Recently, (Jin et al. 2009) compared the accuracy of motion compensation methods for the HRRT. Their study was however limited to comparing their in-house event-based Motion-compensation OSEM List-mode Algorithm for Resolution-recovery (MOLAR), with attenuation and scatter compensated frame-by-frame realignment and post-reconstruction frame-by-frame realignment as well as no motion correction on simulated listmode HRRT data. The latter was built from information from a 120 –minute [11C]-AFM PET scan that is used to study serotonin transporters to create 5-mins long listmode
frames. The motion data input into the listmode data was obtained from 4 real patients, whose movements were recorded using the POLARIS. In the MOLAR algorithm (Carson et al. 2003), the position of each LOR is corrected using the transformations recorded from the POLARIS. The sensitivity image is calculated for every frame so that the effect of motion is accounted for within it. The metrics used for comparison of the motion correction methods were the average intensities of the ROIs, namely raphe nucleus and putamen, and their centroid locations, across the magnitude of motion applied. The authors observed better performance (in terms of the closeness of the measured intensity of the ROI to its true value) using the MOLAR algorithm across all magnitudes of motion as compared to the image registration methods. These were however comparable to MOLAR for movements that were less than 5mm in magnitude but deteriorated with increasing magnitudes. The same observation was noted for the displacement of the centroid location, which served as complementary validation to the ROI analysis as motion could either only slightly shift the position of the ROIs such that the average measure of the ROI would not change. The change in centroid location would then confirm effective movements. Similarly, it is also possible that movements only induce blurring and no shifts of activity in the image leading to a change in average activity but no shifts in centroid location. The reason for the difference in performance between attenuation and scatter compensated FBF realignment and MOLAR is essentially the component of intra-frame motion.

### 3.6 Conclusions

A review of the history of methods developed for motion correction has been presented. It is limited to brain imaging only but covers the key papers in literature for both image-based and event-based methods for motion correction.

Looking at the bigger picture, it can be argued that motion correction algorithms have been developed based on the constraints faced by research groups in terms of technological advancements, data storage and computational power for data processing. The previous generation of PET scanners used to histogram the data online and store
sinograms, which is why, upon the introduction of external motion detection systems like that of (Picard and Thompson 1995) or (Goldstein et al. 1997), on the fly motion correction was deemed a prospect for the future but not computationally feasible at that time. MAFs method was a more viable solution that could be readily implemented in a clinical setting. Modern cameras now offer the flexibility to store every event of the PET data into a listmode data file. This allows for the data to be histogrammed offline, post-acquisition and therefore allows for correction of every event that experienced motion. Since it also allows for temporal flexibility, it offers more control on the ability to define frames *a posteriori* in order to accommodate for within frame motion and still be able to keep *a priori* defined framing protocols.

However, development in data acquisition, storage and reconstruction methods has allowed for motion correction methods to move from image-based to event-based without really assessing the difference between the two based on the ability to reframe the data to avoid intra-frame movements. In comparison assessments by (Fulton et al. 2004; Montgomery et al. 2006; Jin et al. 2009), the scope of assessing the power of reframing the data in a bid to minimise intra-frame movements has been ignored. These authors have conceded that correcting for intra-frame motion would enhance the performance of FBF realignment algorithms. (Costes et al. 2009) has also highlighted this importance in their simulations and tests in an attempt to optimise FBF realignment methods for motion correction. Another problem that is still root for concern in event-based motion correction methods is that it depends on an external measurement system whose accuracy and reliability in a clinical setting is still questioned (Herzog et al. 2005; Rahmim 2005), which is probably why image-based motion correction methods are still popular in PET.

The problem faced by typical clinical research centres like ours is that patients have to be scanned without an external motion tracking system like the POLARIS. This is in turn owing to the fact that despite the system being available, the tracking tool that is fixed to the patient is not well tolerated by this patient group. More generally, there are concerns regarding the tool fixation mechanism used. A neoprene cap, commonly used for this purpose was found to not reflect patients’ head movements accurately. Motion was
visually observed on a substantial amount of those scans and it has been shown by (Montgomery et al. 2006) as well as (Wardak et al. 2010) on how correcting for motion can increase separation between groups and accurately show where regional changes occur between groups.

When motion occurs, especially for discrete motion, often changes can be observed on plots of the total count-rates being detected, conventionally referred to as head-curves, an example of which is shown in Figure 3.3. The head-curve is not likely to be sensitive to all types of motion. The head count is likely to be sensitive to movements that result in the activity within the FOV to change. However, it is possible for some movements like a rotation in the z-direction to not trigger any changes in the head count in that direction. However, as such movements (rotations in the z-direction) are normally effectively prevented by head restraint, the method becomes attractive as it is sufficient to cope with discrete movements that happen in combinations of translations and rotations in any direction. We therefore propose a method for motion detection using centroid analysis which is generated directly from the listmode data. If the head remains still during a scan, the position of the centroid of radioactivity would be constant, except for tracer redistribution, which could be expected to have predictable kinetics, for example, at the early vascular phase, post-injection when the radiotracer moves from the carotid artery to the different tissues of the brain. Using centroid data and taking advantage of the temporal flexibility offered by storing data in listmode, frames can be defined for periods during which there is minimal motion by adding frame partitions at points where motion is suspected. Image-based correction with compensation for attenuation and scatter can then be carried to those newly defined frames as according to (Andersson et al. 1995).

The next chapter examines the benefits of reframing listmode data based on detecting motion in the absence of an external motion tracking tool, as opposed to simply carrying out FBF realignment with compensation for attenuation and scatter and ignoring within frame movements. Motion tracking is effectively achieved by using the raw listmode data itself to generate the centroid of radioactivity, which is the average activity over the
scanner’s FOV, across the duration of the scan and relating the changes in the centroid to movements observed during the scan.

**Figure 3.3:** Plot of the head curve (counts per second) versus time for randoms (red), prompts (green) and trues (blue) for an $^{18}$F-FDG PET scan. Discrepancies can be observed on the curves between time intervals 900-1200; and 2100-2400 for example, that could be due to motion.
Chapter 4
Motion Correction using improved frame-by-frame realignment based on reframing listmode data using centroid analysis

4.0 Introduction

One conclusion from chapter 3 is that event-based methods for motion correction are better in performance than image-based ones largely due to the fact that intra-frame motions are ignored in the latter.

In general, methods for motion correction that require external detection and quantification of the movements have been found to be undesirable clinically, with questionable accuracy and reliability of the quantification of the detected movements.

The practicality of acquiring PET data in listmode format has also been appreciated in its ability to reframe data post acquisition. The ability of using the head-counts as an indicator to when movements occur has also been speculated on.

From these observations, the following questions are formulated:

1. Can the raw listmode data be used to detect when motion occurs?
2. In practice, does the reframing of the listmode data, according to when movement occurs, result in reduced movement degradation of the data?

In this project, we address the problem of intra-frame motion in frame-by-frame realignment by taking advantage of the temporal flexibility offered by acquiring and storing data in listmode format. This chapter outlines the proposed method for improved frame-by-frame realignment, using centroid analysis as a method for detecting movements. It assesses the use of centroid analysis as a tool for motion detection and evaluates the added benefit of reframing emission data prior to frame-by-frame
realignment against frame-by-frame realignment algorithms that are popularly used in clinical and research settings.

4.1 The Centroid of Activity

4.1.1 Method for calculation of the centroid position

The centroid of radioactivity within the camera field of view (FOV) is the average position of that radioactivity and is mathematically given by the product of the amount of radioactivity and its position integrated over the FOV. Its position can be determined from reconstructed images, but such an approach would require image reconstruction and would have limited temporal resolution. Examining the lines of response between two opposing heads of the HRRT, as shown in Figure 4.1, a “rectilinear” type 2D image can be formed, from which the centroid of activity within this projection can be determined.

The rectilinear images shown in Figure 4.1 were calculated using an incidence angle width of 26. The incidence angle width is an offset that is applied to every detector crystal in order to allow LORs coming from a limited width. Varying the offset angle is synonymous to finding a trade-off between resolution and Signal-to-Noise Ratio (SNR). As the offset difference is increased, there is more information read over that plane. However, there is a loss in activity towards the edge of the head planes being constructed. The resolution of the image is therefore likely to degrade but SNR would increase.
Figure 4.1: Layout of the HRRT with its 8 octagonal heads. By considering coincidence events over ‘Top’ and ‘Bottom’ heads (shown in blue), 2D rectilinear images, and hence centroid coordinates can be obtained for the x-axis and the z-axis as shown. Similarly, by considering coincident events across the ‘Left’ and ‘Right’ panels of the heads (shown in red), rectilinear images are obtained for the y-axis and z-axis. Integration of the product of the activity and position over the FOV gives the centroid. (*) shows example of the incidence angle width set here as a maximum. The incidence angle is an offset angle that describes the maximum allowed offset that accepts those coincidence events.

From the rectilinear images, the centroid of activity between the top and bottom heads during a short time period can be defined as equations 4.1 and 4.2 respectively:

\[
C_x = \frac{1}{N_S} \sum_{i \in S} \frac{x_{i,1} + x_{i,2}}{2} \quad \text{Equation 4.1}
\]

\[
C_z = \frac{1}{N_S} \sum_{i \in S} \frac{z_{i,1} + z_{i,2}}{2} \quad \text{Equation 4.2}
\]

where

\(C_x\) and \(C_z\) are the positions of the centroid of activity along the x-axis (left to right) and the z-axis (superior to inferior) respectively;
The positions of the crystals detecting a coincident event along the x-axis and z-axis respectively are $x_i, \ y_i$ and $z_i, \ y_i$.

$S$ are the set of coincident events which are detected between the top and bottom head and with an incidence angle within specified offset limits and within a specified time period; and

$N$, are the number of coincident events within the set $S$.

Similarly, using the left and right heads, the centroid positions can be estimated along the y-axis (anterior to posterior) and a second estimate of the z-axis can be obtained. The 2 values for the z-axis provide an additional benefit of checking for data consistency.

### 4.1.2 Performance of the centroid method

A series of experiments were carried out to determine the performance of the centroid in its ability to deliver a sensible estimate of movements that has occurred during a scan. A cylindrical phantom, 140 mm in diameter, 120 mm long, was filled with 25 MBq of $[^{18}\text{F}]$-FDG and positioned rigidly in the headrest of the bed of the HRRT. Scan data was acquired by moving the phantom using the scanner bed, by producing discrete translations in the y-direction (up-down) and the z-direction (head-feet) and scanning each position for 300 seconds.

Figure 4.2 (column a) shows the 2D rectilinear images for the cylindrical phantom over the first 300 seconds of the scan, showing the distribution of the position of coincidence events in an environment of no movements between top and bottom heads (X-Z plane, top row of Figure 4.2) and between left and right heads (Y-Z plane, bottom row of Figure 4.2). It can be easily observed from both top and bottom figures of figure 4.2a that the centroid tends to be biased towards the centre point of the plane than what is normally expected off a uniformly filled phantom where the centroid is expected to be along the central axis of the phantom. This is because the incident angle between the opposing heads being considered is set to a maximum, such that, there is a loss in activity towards the edge of the head planes that leads to a biasing of the centroid position from its true
position to the centre point within that plane. Additionally, due to the limited FOV, distortions are observed towards the edge of the object.

**Figure 4.2:** 2D rectilinear images across the X-Z planes (top row) and Y-Z planes (bottom row) showing the distribution of the position of the coincident events for the first 300 seconds of scanning the cylindrical phantom, calculated and presented in descending order of incidence angle between opposing heads with (a) incidence angle width = 104 (maximum i.e. set S all counts between heads), (e) incidence angle width = 1 (minimum). As we go from (a) to (e), centroid becomes more accurate but there is a loss in SNR implying that detecting movements gets more difficult as the data is more noisy.

### 4.1.3 Accuracy and sensitivity versus Signal to Noise Ratio of the centroid

By restricting the angle of incidence, the distribution of coincident events becomes more ‘even’ such that the eventual calculation of the centroid of activity is more accurate, as shown in Figures 4.2b and 4.2c respectively. However, the increase in accuracy also brings along a decrease in the Signal to Noise Ratio (SNR) of the distribution of coincident events as there are effectively less counts used for the calculation of the centroid (in this case, ~99 kcounts/s for maximum incidence versus ~63.9 kcounts/s for a width of 52 versus ~19kcounts/s for a width of 26 versus 1.8 kcounts/s for a width of 8 and versus 8 counts/s for a width of 1). Figure 4.3 shows the time course of the centroid for the cylindrical phantom sampled at every second over the total duration of the scan. It can be observed that while the sensitivity and accuracy of the centroid is higher with a smaller incidence angle between the opposing detector heads, the SNR is better when the incidence angle is set to maximum. This implies that while sacrificing the accuracy of the centroid
measurement, it is output with a better precision, such that changes in centroid measurements can be more easily related to movements (especially in the case of discrete movements).

**Figure 4.3:** Graphs showing the time course of the centroid across the Z-axis over the total duration of the scan for the cylindrical phantom. **Top:** Centroid sampled at every second calculated with different incidence angles between opposing heads with width = 104 = maximum (blue), width = 52 (green), width = 26 (red) and width = 8 (yellow). It is observed that the yellow graph is noisier than the red one which is in turn noisier than the green one and that the blue is the least noisy. The value for the centroid for each graph is different, confirming that accuracy is affected by the incidence angle chosen. Since the yellow graph is so
noisy that the other centroid time courses are barely visible, the same graph is plotted again at the bottom to show centroid for widths 26, 52 and 104. Note: centroid for width=1 was not plotted as it was too noisy and spread itself over the whole of the graph.

The bias between centroid measurements is not linear as observed from Figure 4.3 above. Between t=0 and t=300 seconds, it averages crystal position 44.2 for width 26, 45.4 for width 52 and 47 for maximum width; while for between t=750s and t=1000s, it averages crystal position 60.7 for width 26, 59.7 for width 52 and 57.7 for maximum width. One explanation for this is that as the bed position moves, it might partially step out of the FOV such that there are even less coincident events available for calculation of the centroid. Since the number of coincidence events going in or out of the FOV is non-linear, the difference in centroid measurements at different bed positions for different incidence angles will be equally non-linear. From an object moving out of the FOV, it can also be inferred that due to this movement, bias in the centroid will still be recorded despite a reduction in offset/incidence angle.

4.1.4 Extension of the observation to patient’s data

Within patients’ data, the importance of a higher SNR becomes more obvious as shown in Figure 4.4. The red graph is so noisy that it is difficult to distinguish between noise and movement of the centroid (in the regions encircled). The situation is improved on the green graph which has a higher incidence angle. However, differentiating between noise and movements is potentially still problematic as shown in the encircled regions. For all the encircled regions shown in Figure 4.4, the blue graph corresponding to the centroid calculated with maximum incidence angle, shows the highest precision and therefore increases the confidence in finding the point when movement effectively occurs.

It is quite clear from the cylindrical phantom that centroid calculations as a measure of movements would be insufficient as they tend to be insensitive to rotational movements. Moreover, bias is always likely to result if the subject moves out of the FOV, even if small offset angles are used. Since the aim here is to detect when motion has occurred as opposed to quantifying the motion that occurred, precision takes priority over accuracy.
For that reason, a maximum incidence angle was kept for calculation of centroid of activity in patients’ data.

![Graph showing centroid data](image)

**Figure 4.4:** Centroid data for a patient (scanned with $^{18}$F-FDG across the Z-axis over the total duration of the scan. Centroid calculated with different widths of incidence angles between opposing heads with width = maximum (blue), width = 52 (green) and width = 26 (red). Encircled are the regions where it is difficult to distinguish between movements of the centroid or noise. The blue graph presents the centroid data with the highest SNR which increases the confidence in distinguishing between movements and noise. Note: widths of 8 and 1 not shown because data is too noisy for a proper illustration.

### 4.1.5 Quantitative implications of sacrificing the accuracy of the centroid

To quantify the loss in accuracy of the centroid in order to obtain better SNR, the emission data for the cylindrical phantom was reconstructed without attenuation and scatter. For each frame, the position of the centroid on the image was calculated from the listmode data and marked on the emission image. Each frame was also co-registered to the baseline initial image acquired in order to quantify motion from the transformation parameters obtained off the co-registration process. Since a physical phantom is not affected by tracer redistribution, the discrepancy between centroid positions calculated from the listmode data versus that obtained from co-registration parameters would reflect the loss in accuracy due to using a different incidence angle.
Prior to comparing the translations obtained from the centroid to that using co-registration parameters, the robustness of the co-registration method in delivering accurate measures of translations was confirmed by comparing the co-registration parameters with the movements recorded from the coordinates of the bed position as shown in Table 4.1 below. It can be observed that the co-registration parameters recorded were in-line with the bed transformations applied to simulate movements.

<table>
<thead>
<tr>
<th>Bed positions (mm)</th>
<th>Bed transformations (mm)</th>
<th>Co-reg parameters (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td>273</td>
<td>496</td>
<td>0</td>
</tr>
<tr>
<td>273</td>
<td>493</td>
<td>0</td>
</tr>
<tr>
<td>273</td>
<td>485</td>
<td>0</td>
</tr>
<tr>
<td>288</td>
<td>480</td>
<td>15</td>
</tr>
<tr>
<td>303</td>
<td>480</td>
<td>30</td>
</tr>
<tr>
<td>303</td>
<td>470</td>
<td>30</td>
</tr>
</tbody>
</table>

**Table 4.1:** Bed positions recorded from the scanner and using the first position as reference, the subsequent transformations (translations only) of the bed as well as those translations determined from reconstruction of the phantom without attenuation and scatter and co-registering to the reference position. It can be seen that the transformations obtained from co-registration agree with the translations applied to the phantom via moving the bed.
**Figure 4.5:** Relationship between centroid positions measured from listmode data along the Z-axis versus as measured from co-registration transformation parameters for different incidence angles of widths 26 (red), 52 (green), and maximum (blue). Choosing a width of 52 gives a relationship that is close to 1-1, a width of 26 over-estimates the measurements while the maximum incidence angle tends to under-estimate it.

The graphs in Figure 4.5 indicate the correlation between centroid positions calculated from the listmode data versus calculated from co-registration parameters. The co-registration parameters are the same ones recorded in Table 4.1. The measured listmode is obtained by calculating the median value of the position of the centroid obtained over that period of time. The centroid positions can be seen in Figure 4.3. Predictably among the 3 different incidence angles tested, the centroid offers the most accurate measurement for translation across the z-axis when the width was set to 52 (gradient =1.04) as it offers a good cut-off for the number of coincident events detected. Choosing a too low incidence angle tends to over-estimate the movements (gradient =1.21) and as seen in figures 4.2 – 4.4, offers a noisier distribution of the centroid values. When the incidence angle is set to maximum, the position of the centroid is under-estimated (gradient =0.72) as a higher number of random events are included in its calculation. However, as observed in Figures 4.2-4.4, it offers a higher SNR that makes detection of centroid changes due to movement easier. Figure 4.4, which is patient-data-oriented, also
strengthens this reasoning. Figure 4.5 indicates that if the same comparisons were to be extended to patients’ data, a gradient of 0.7-0.8 would indicate that the method is robust in detecting movements that has occurred during the scan. A gradient of higher than 1 or below the 0.7 mark is more likely to indicate that the centroid is hampered by tracer redistribution over the duration of the scan and/or that the co-registration of the emission frames are not accurate enough.

4.1.6 Conclusions from physical phantom experiments regarding use of centroid analysis:

The main conclusions that can be drawn from the above experiments are:

1. The centroid is not likely to be sensitive when only rotational movements are observed. It therefore cannot serve as a method for quantifying movements.
2. Bias in the calculated centroid is likely to result irrespective of the chosen incidence angle as the subject/object moves out of the FOV.
3. Because of the above two points, the centroid has been developed mainly to serve as a method to detect for motion rather than quantify motion.
4. Since precision rather than accuracy is critical in this case, it is best to keep the incidence angle between opposing detectors to a maximum when calculating the centroid of activity during a scan as it increases SNR and makes centroid changes due to movements easier to detect.
5. It is a potentially serious limitation of the method that the centroid is not likely to be respond when there are changes due to rotations only. However, from our dataset of patients suffering from dementia, it was observed by clinicians at the WMIC that most movements that occurred were fairly large discrete movements where rotational movements normally occurred in conjunction with translational movements. These are known to create gross misalignment of the position of the head between acquisition of the emission and transmission data. It was therefore believed that using the centroid to locate those mismatches and correct for them was worth doing.
We therefore use the centroid as a tool for motion detection, maximising on temporal resolution by analysing the centroid at every 1 second interval and on the signal to noise by keeping the incidence angle at a maximum. We have designed and implemented an interactive tool for visualisation and interpretation of centroid data to eventually relate the changes in centroid coordinates in the X, Y and Z axes to movements of the patients during a scan. Based on those movements observed, the frames are redefined using the GUI to give new frames where the movements that occur within formerly predefined frames have been minimised.

4.2 Reframing the listmode data based on output of the centroid

4.2.1 Method for analysing the centroid output to reframe the listmode data

Figure 4.6 shows an example of the centroid position during a scan where the patient received 370 MBq of $[^{18}\text{F}]$-FDG via intravenous injection and scanned on the HRRT for 60 minutes. The scan was preceded by the acquisition of a transmission scan of 6 minutes that is used for deriving the attenuation and scatter correction factors during the reconstruction of the emission data using OP-OSEM iterative reconstruction method, 16 subsets, and 10 iterations. However, this transmission time of 6 minutes should not be confused with the waiting time prior to dose injection (445 seconds in this case). The waiting time of 445 seconds follows the acquisition of the transmission data.
Figure 4.6: Example of the time course and position of the centroid for a real patient for an [18F]-FDG scan along the scanner’s x-axis (C_x in blue), y-axis (C_y in green) and the z-axes (C_z1 and C_z2 red and yellow). Within the figure:

1. Start of injection; 2. A-priori defined frame of 70 seconds post-injection to observe the vascular phase of the tracer; 3. A-priori defined frame at 1200 seconds post injection which is a transient frame that looks into accumulation of the tracer in different regions of the brain; 4. A-priori defined frame of 2400 seconds which is defined as the ‘static’ frame that is used at the WMIC for further data analysis and represents glucose metabolism in the different regions of the brain. Description of movements is as follows: P denotes step changes in the centroid values and is subscripted to their relative axes. They relate to discrete movements. Q denotes slow changes in the centroid values across time in the respective axes. Whether the change relates to motion or tracer redistribution is not fully known. The data is reframed in a conservative way such that it allows for consideration for motion over such periods. R denotes the early vascular phase of the tracer where the first 2 frames can be concatenated as the change in centroid at that point is more likely to be due to tracer redistribution than movements. The idea is to allow the user to insert a new frame partition (shown as the dotted vertical lines) into the pre-defined schema (shown as the solid vertical lines), so that only small changes in the centroid position occurs within each frame.

Figure 4.6 illustrates the various types of shifts in centroid position that can occur. They can be separated either into discrete step changes in centroid coordinates or into drifts where the centroid coordinates slowly slopes upwards or downwards over a considerable period of time. While the step changes (P_x, P_y) can be attributed to discrete motion, the progressive continuous motion could be related to either the rapid distribution of activity into the brain within the first 1-2 minutes following injection (R_x) or a slow progressive
head movement \((Q_y, Q_z)\), of the patient for example, when the latter is falling off to sleep. The idea is to allow the user to insert a new frame partition (shown as the dotted vertical lines) into the pre-defined schema (shown as the solid vertical lines), so that only small changes in the centroid position occurs within each frame. For discrete step changes, choosing such a position is fairly straightforward as the point where the step change occurs is easily located on the graph. However, in conditions of continuous motion of the centroid, it is sometimes difficult to discriminate between tracer redistribution and continuous movement of the patient. A more conservative approach would then be to partition the frames at a reasonable frequency to accommodate either cause.

In the reframed data shown above, for the first two frames of in the time-course, it can be argued that early centroid changes denoted by \(R_z\) are essentially due to tracer redistribution than actual movements. The interactive tool has been designed so that 2 successive frames can be assumed to have no motion in between them and can be concatenated together before reconstruction without attenuation and scatter. This allows the resulting emission frame to have better statistics for a more accurate co-registration.

The output of the centroid analysis from the interactive tool is a new frame definition that has the following desired characteristics:

1. It contains very little motion within frames, reflected by the fact that the position of the centroid does not change significantly within each frame.
2. The former a-priori defined frames are still part of the new schema. These frames can be concatenated post-reconstruction to revert back to the old schema.

### 4.3 Proposed reframed-realignment algorithm

An improved frame-by-frame realignment algorithm is proposed that stems from the works by (Andersson et al. 1995) to allow reframing of the listmode data to account for motion based on analysing the centroid position over the duration of the scan, as defined in section 4.2, and have aptly named it Reframed-Attenuation-Scatter compensated Frame-By-Frame (RAS_FBF) realignment. The 7 main steps of RAS_FBF are described below:
Step 1: New frame definitions

The frame definitions are refined using centroid analysis as described in section 4.2.

Step 2: Reconstruction of motion uncorrected images

Using the modified frame definitions, images are reconstructed using OP-OSEM, 4 iterations, 16 subsets, in their respective native space without correction for attenuation and scatter, using standard normalisation and correction for randoms. The term ‘native space’ refers to the actual position of the head in the scanner. Reconstructing without attenuation and scatter corrections prevents the induction of errors due to these parameters that work only if the emission data is aligned with the transmission data. Importantly, these reconstructions provide an increased signal in the emission images that ultimately increases the accuracy of the registration process (Costes et al. 2009).

Step 3: Co-registration of emission images to an image of attenuation coefficients (µ-map)

From the transmission scan, an image of the attenuation coefficients is reconstructed using the HRRT Users’ Community software which reconstructs a µ-map using the routine umap_process_u(Sibomana et al. 2004). The registration of each reconstructed emission frame to this µ-map is done in 2 stages:

In the first stage (3a), a ‘reference frame’ emission image (EM_{TX}) is chosen with good statistics and minimal motion within the frame, typically from the middle of the scan to avoid tracer redistribution effects and maximise on the quality of EM_{TX} for robust co-registration with the reconstructed emission images. This reference frame is registered to the µ-map image via rigid-body registration. The use of a reference frame instead of the transmission image to co-register the reconstructed emission images is desirable as registration with the µ-map is challenging due to lack of anatomical information within. This, coupled with the fact that some of the emission images are potentially noisy, contributes to affecting the robustness of the registration process if the µ-map image is used. When co-registering the µ-map image to the reference emission image (EM_{TX}), Normalised Mutual Information (NMI) is used as criteria for optimising the goodness of fit.
between those images as the characteristics of the µ-map image and the emission image differ (Maes et al. 1997; Maes et al. 1999).

In the second stage (3b), each emission image is registered to the reference frame EM\textsubscript{TX} and thereby registered to the µ-map image. Since the emission images are from the same modality, correlation coefficient was the preferred similarity measure used to optimise the registration process.

The image co-registration steps are performed in VINCI (Cízek et al. 2004). The output of stage 3b are parameter files in xml format that contain the 6 degrees of freedom that map the transformation of each emission image to the reference µ-map image space. Most importantly, it allows easy calculation of the inverse of the transformation that is uniquely defined so that the µ-map image can be mapped onto every emission image.

**Step 4: Creation of motion compensated attenuation and scatter correction sinograms**

Using those inverse transformations, the µ-maps are resliced to the ‘native space’ of each emission frame. The attenuation correction sinograms are derived from the integration of the registered µ-maps along LORs and taking their exponentials (Nuyts et al. 1999). Scatter correction is performed for each of the defined emission frames using the corresponding resliced µ-maps and measured trues for every defined frame using Single Scatter Simulations (Watson 1999). The attenuation and scatter correction sinograms were created using standard HRRT software routines create\_3d\_atn\_u and scatter\_process\_u respectively (Sibomana et al. 2004).

**Step 5: Reconstruction with appropriate correction factors**

The emission images were then reconstructed with OP-OSEM, 10 iterations, 16 subsets in their ‘native space’, using their respective frame specific attenuation and scatter correction factors, and applying standard normalisation procedures as well as correction for randoms.
**Step 6: Reslice to initial reference space**

Using the rigid-body transformations defined in Step 3, each reconstructed emission frame image is resliced back to the position of the brain during the transmission scan.

**Step 7: Frame concatenation**

The frames that were subdivided to accommodate for motion are concatenated to revert back to the original frame definitions. Prior to that, they need to be calibrated (corrected for frame duration, dead time, branching factor etc). For the concatenation process, the final motion corrected frame is the weighted average of the number of frames that it was divided into.

The steps described above are summarised in a flowchart in Figure 4.7 (left hand side) further below.

### 4.4 Performance assessment of RAS_FBF method

The benefits of reframing emission data to minimise intra-frame motion detected using centroid analysis was assessed by comparing the performance of RAS_FBF against:

(a) No correction for movements

(b) FBF – conventional method for frame-by-frame realignment, post-reconstruction, as outlined by previous authors in Chapter 3. (Woods 1992; Mawlawi et al. 2001)

(c) AS_FBF – Attenuation and Scatter compensated Frame-By-Frame realignment method that allows realigning of the \( \mu \)-map to the ‘native space’ of every emission frame in order to achieve more accurate corrections for photon attenuation and scatter as described above. The only difference between RAS_FBF and AS_FBF is that the frame definitions in AS_FBF do not change during the motion correction process while in RAS_FBF, the emission data is reframed according to where motion has been detected using centroid analysis. (Andersson et al. 1995)

Figure 4.7 shows a flowchart that summarises and contrasts the steps involved in FBF, AS_FBF and RAS_FBF. FBF realignment consists of choosing a reference emission frame post-reconstruction and co-registering the remaining frames to that reference frame. The
method overlooks any mismatch between emission and transmission images. In AS_FBF, the emission data is compensated for attenuation and scatter by only applying Steps 2 – 6 on the acquired listmode data and keeping the original frame definitions at all times. Neither FBF nor AS_FBF has the added feature of motion detection or make use of the temporal flexibility offered by listmode data acquisition and storage for reframing the emission data.

4.4.1 Materials for evaluation of the proposed RAS_FBF method

We used 6 PET scans from a clinical study involving early AD patients, where we are looking at glucose metabolism in the brain following intravenous injection of 370 MBq of $^{18}\text{F}$-FDG and scanning for 60 minutes on the HRRT. For each scan, 2 frame definitions were reconstructed with OP-OSEM, 16 subsets, 10 iterations. The first frame definition is a ‘static’ frame of 2400s that is acquired at 1200s post-injection, and is used for qualitative analysis. The second frame definition consists of 33 frames post-injection with timings 1x10s, 12x5s, 2x10s, 3x30s, 3x60s, 2x120s, and 10x300s that are used for quantitative analysis of tissue kinetics.
Figure 4.7: Flowcharts for the proposed method (RAS_FBF) that allows for reframing data to suppress motion within frames as well as compensate for attenuation and scatter corrections; AS_FBF which provides motion compensated attenuation and scatter correction to images but no reframing; and FBF which is the conventional frame-by-frame realignment applied post-reconstruction in most clinical studies.
4.4.2 Methods for performance evaluation of the proposed method.

4.4.2.1 Comparison of centroid and registration detected motion

Prior to looking into the performance of the proposed method as a whole, we use the centroid data to make a comparison of the motion measured from the position of the centroid as calculated from the listmode data versus its position as calculated from the transformation parameters output from the co-registration step. The method is similar to that applied to the uniform cylindrical phantom in section 4.1.5. For the 6 FDG-PET scans, for each frame defined from the centroid analysis, the position of the centroid was determined from the listmode data. To determine the corresponding position of the centroid of radioactivity for every emission image (in ‘native space’) from the registration process, the rigid-body transformations that describe the movement of the reference frame to the ‘native space’ of the brain was applied to the position of the centroid in the reference frame. This resulted in obtaining the position of the centroid only due to motion of the brain in the scanner’s FOV and not because of how the centroid of radioactivity changes with tracer redistribution or with the position of the brain within the scanner’s FOV. This resulting centroid position was compared with that calculated straight from the listmode data. The position of the centroid coordinates were calculated in millimetres, with the centre of the scanner’s FOV used as the origin.

4.4.2.2 Qualitative Assessment of motion correction methods

We used frames of the mean radioactivity concentration from 20-60 minutes post injection from the 6 $[^{18}\text{F}]$-FDG PET scans reconstructed with (i) No motion correction, (ii) FBF, (iii) AS_FBF, and (iv) RAS_FBF. The data was presented in a blinded and random fashion to 4 observers (all clinicians who routinely analyse PET scans at the WMIC) to evaluate quality of the images. They rated image quality on a scale of 1 to 5, in ascending order, and were requested to briefly comment on the features they observed within the images to decide upon their ratings.
**4.4.2.3 Assessment of motion correction methods using Time-Activity Curves**

Time-Activity Curves (TACs) were generated from the 6 FDG patient scans using dynamic frame definitions over 60 minutes following injection and compared between the four methods (no movement correction, FBF, AS_FBF, and RAS_FBF). The criterion used for evaluation was predictability of tracer accumulation, based on the continuous physiological accumulation of FDG. We looked at TACs for the left and right rectus gyrus of the brain as it is positioned quite far away from the typical axis of head rotation, which lies in the upper neck region, and therefore experiences significant displacements associated with small rotational movements. Both the left and right region were analysed separately as an added assessment, since their metabolism is expected to be the relatively similar. However, we only report the observation as results are shown for the left rectus gyrus only. The right rectus gyrus was omitted in the report as it was found to be difficult to appreciate the differences between motion correction methods when more than one region was plotted, especially when their metabolisms are similar. Plotting the result for the right rectus gyrus separately would look like reporting the same result twice and not be of any direct benefit towards making the point of how the different motion correction algorithms performed.

**4.4.3 Results**

**4.4.3.1 Comparison of centroid and registration detected motion**

When comparing listmode- and registration-derived movements in the 6 FDG PET scans, it was observed that displacement of the centroid was predominantly in the axial z-direction, with only modest displacements, of magnitudes less than 0.5 mm, observed in the transaxial x- and y-directions for the majority of the scans. However, out of the 6, one scan did show significant displacements of the centroid in the x-direction and is therefore reported in Figure 4.8a. Displacements across the z-axis for all 6 scans are reported in Figure 4.8b. Comparing the listmode and registration derived centroid positions it was observed that:

i. During early frames, defined as those frames that lie within the 5 minutes following injection of $[^{18}\text{F}]$-FDG, a proportional relationship between listmode and
registration based derivation of centroid positions was less clear. One possible explanation is tracer redistribution that occurs at the early vascular phase when the tracer moves from the carotid arteries into the various parts of the brain. This gets accounted for in the listmode-derived calculations and shifts the centroid to a more superior position in the scanner’s FOV. Another explanation is that short frames suffer from low statistics and are inherently noisy such that the subsequent image registration process is prone to inaccuracies.

ii. For the later frames, a proportional relationship similar to that of the phantom data was observed for both, the x-axis (gradient = 0.66) and the z-axis (gradient = 0.88).

**Figure 4.8:** Comparison of centroid positions as measured from listmode data versus measured from transformation parameters from the registration process. Centroid positions for the early frames (defined as frames within the first 5 minutes following injection) are shown in red. Later frames are shown in blue. (a) Only one scan showed significant movements in the x-axis and is reported. (b) centroid position data for all 6 scans in the z-axis.
4.4.3.2 Qualitative Assessment

Figure 4.9: Graph showing the mean ratings and their standard deviations for 4 independent observers and uncorrected data and 3 different methods used for motion correction. Description of each method is summarised in Figure 5.7

Figure 4.9 shows the mean ratings over the 4 observers and their standard deviations for each motion correction method. The proposed method of reframing using the centroid analysis, RAS_FBF, scored the highest average rating between observers (3.75 ± 0.68), followed by AS_FBF (3.21 ± 0.59), then FBF (2.33 ± 0.48). The uncorrected data always scored worst (1.25 ± 0.44), with the observers claiming the data to be blurred, noisy in appearance and difficult to delineate anatomical regions like the cortical gyri. When comparing with other methods, the observers further looked at improvements in intricate regions of the anatomy that are characteristically visible on an HRRT image, like the separation between the caudate and the putamen, or how well resolved the colliculi are. An Analysis of Variance (ANOVA, Tukey test) for multiple comparisons with rating between observers as dependent variable showed significant differences between methods (p<0.001). For comparison between RAS_FBF and AS_FBF, significant differences were found (p=0.008). No significant changes were found for ratings between observers (p=0.77) despite the observers being blind to the
method of correction used indicating consistency in ratings between observers and low variability in subjective assessment. This would in turn probably mean that the changes observed in ratings were only due to the method of motion correction used.

In addition to the statistical analysis it was noted that for every dataset, all observers rated the correction methods as RAS_FBF >= AS_FBF >= FBF >= no correction. Figure 4.10 shows two examples that were corrected with the 3 methods. Both, Subject1 and Subject2 show improvement between uncorrected data and FBF, with a better delineation of the brain anatomy, especially along the cortical gyri. The further improvement in contrast to noise ratio is also visible when comparing FBF to AS_FBF in both datasets. However, for Subject1, RAS_FBF received higher scores than AS_FBF having a brighter cortical pattern, whereas for Subject2 similar scores were received.

4.4.3.3 Assessment of motion correction methods using Time-Activity Curves

Time Activity Curves (TACs) for a region over the left rectus gyrus for Subject1 and Subject2 and for the 3 motion correction methods plus no correction are shown in Figure 4.11.

Clear steps in the TACs for both subjects are observed with no correction. In Subject1, there is an apparent reduction in the accumulation of FDG with FBF over the first 3000s. AS_FBF for Subject1 shows a higher uptake over the total duration of the scan as compared to FBF but is never higher than the uncorrected data. RAS_FBF has a higher uptake than AS_FBF over the duration of the scan but step changes between AS_FBF and RAS_FBF is absent. Both AS_FBF and RAS_FBF result in a better steady increase in activity concentration as opposed to the uncorrected and FBF data.

In Subject2, there is an apparent increase in uptake between the uncorrected data and FBF over the first 2500 seconds of the scan but no apparent difference between FBF and AS_FBF over that same period. RAS_FBF shows a slightly higher uptake over that period.

The rectus gyrus was chosen because it is a particularly metabolically active metabolic region and prone to movements such that motion, including within-frame motion, is likely
to lower the radioactivity concentrations. The TACs tend to reflect the observation that radioactivity concentrations obtained via each method ranked with the qualitative assessment in that the uncorrected data, FBF, AS_FBF and RAS_FBF rank themselves in this ascending order of quality and hence performance.

Figure 4.6 is in fact the centroid data for the motion corrected Subject1 and it can be observed that what is believed to be discrete motion, from the step changes in the centroid coordinates between Time=2500s and Time=3500s, actually translates to step changes in the TACs over that same time period. The TACs of Subject1 also confirms the visual findings that there is an increase in activity in brain regions when corrected with RAS_FBF than with AS_FBF. This is in line with previous works (Jin et al. 2010) which is discussed further below.

While it is difficult to observe the benefits of reframing the data when visually assessing the quality of the images for Subject2, quantitatively, higher time activity curves are seen; between Time=3000s and 3500s, RAS_FBF seems to output a more steady increase in activity concentrations than AS_FBF over that period.
Figure 4.10: Examples of motion corrected data using the specified methods with observer ratings. Subject 1 shows marked qualitative improvement as the cortical delineations look brighter in RAS_FBF than in AS_FBF. For Subject 2, distinguishing between AS_FBF and RAS_FBF is more difficult.
Figure 4.11: Time-Activity Curves of the left rectus gyrus for Subject1 and Subject2 for the uncorrected data (black), data motion corrected with: FBF (green), AS_FBF(blue), RAS_FBF(red). Subject1 confirms visual findings that the activity within the brain regions increases from AS_FBF to RAS_FBF. In Subject2, despite no obvious differences between AS_FBF and RAS_FBF, the latter seems to have a more steady increase in tracer uptake than AS_FBF between Time=3000s and Time=3500s.

4.5 Discussion

4.5.1 Impact of the work and key findings

In chapter 3, the pros and cons of event-based versus image-based reconstruction were discussed. The key findings that have prompted the exploration of the proposed method (RAS_FBF) as a means for motion correction are as follows:

1. Event-based is superior to image-based motion correction (Bloomfield et al. 2003; Fulton et al. 2004; Montgomery et al. 2004) as it was found to yield better quantification. One of the main reasons for this is that event-based motion correction allows for correction of every LORs detected during the scan hence maximising on temporal resolution by correcting for motion within frames. Yet, the use of frame-by-frame realignment methods for motion correction, and finding ways for improving it (Costes et al. 2009; Jin et al. 2010; Wardak et al. 2010), in a clinical environment still
has utility as image-based motion correction is computationally faster than correcting LORs and allows for retrospective movement correction in the absence of an external system for motion detection. It has also been highlighted that for event-based methods to be accurate, the external instrumentation system used for detecting and quantifying the movements must be equally accurate and reliable. (Herzog et al. 2005; Rahmim 2005).

2. Another obvious reason is that the popular frame-by-frame realignment methods in literature (Woods 1992; Andersson et al. 1995), against which event-based methods have been compared, do not offer the possibility for correcting for movements that occur within a-priori defined frames. In fact, (Montgomery et al. 2006) does concede that the performance of their frame-by-frame realignment method would be similar to that of their event-based correction method if there were no movements within frames.

3. Methods developed towards improving frame-by-frame algorithms have been geared towards improving on the co-registration between transmission and emission images (Costes et al. 2009) or segmenting the head rest out from transmission image to derive more accurate μ-maps (Wardak et al. 2010). To my knowledge, there have been no works carried out on the possibility, and evaluating the impact of reframing the emission data prior to motion correction via frame-by-frame realignment.

With data acquisition and storage in listmode format now more and more routine in clinical practice, we have exploited its temporal flexibility in 2 ways. Firstly, in calculating the centroid of activity by sampling the listmode data for coincident events and their positions at every second interval. Secondly, by tracking the movement of the centroid over the duration of the scan, we locate changes in centroid positions and reframe the data to produce frames within which there are minimal changes in centroid coordinates, with the aim that this corresponds to minimal movements within frames. We then evaluated the added benefit of reframing the data by comparing the proposed method (RAS_FBF) against conventional frame-by-frame realignment (FBF) as according to (Woods et al. 1992) and implemented in (Mawlawi et al. 2001), as well as against attenuation and scatter compensated frame-by-frame realignment (AS_FBF) as designed and implemented by (Andersson et al. 1995).
Qualitative indices from the results section have shown that RAS_FBF performs best, followed by AS_FBF. AS_FBF itself performs better than FBF and it was observed that FBF still provided qualitative improvement over no motion correction. These findings corroborate with those from other authors (Fulton et al. 2004; Montgomery et al. 2004) regarding the improvement brought to image quality with motion correction and further improvement based on how the frame-by-frame algorithm is refined. However, it has to be noted that qualitative changes between AS_FBF and RAS_FBF are dependent on the presence of within frame motion. If there is no significant within frame motion, AS_FBF and RAS_FBF will be the same. (Jin et al. 2010) reported that motion could impact on images in 3 ways: it would either cause blurring of the region, or distortion of the region, or both. If the impact of movements is blurring of a region, assessing its qualitative improvement, especially between AS_FBF and RAS_FBF, is potentially difficult as the emission frames are subjected to rigid-body transformations that are followed by trilinear interpolation during re-orientation back to the reference space. The important note here is that benefits of reframing the data can sometimes be missed through a qualitative analysis.

However, these become clear in analysis of TACs where the shift of activity within a region (example of the left rectus gyrus) is observed in TACs through step changes between frames for the uncorrected data. These step changes are reduced in magnitude as the correction methods improve from none to FBF to AS_FBF and to RAS_FBF. Also, in general, the TACs for RAS_FBF show that the activity within the region of interest goes highest as opposed to AS_FBF or FBF. The best explanation for this observation is that we are looking at essentially grey matter regions that contain more activity than the surrounding white matter regions. Motion results in blurring which for a hot ROI means a decrease in activity concentrations or a shift in activity concentrations that results in the same effect. Correction for movement between successive frames via FBF or AS_FBF improves the activity measurements in the ROI. Correction for movement within frames results in further improvements that are seen using RAS_FBF.
4.5.2 Limitations of the method

In order to increase the SNR of the centroid analysis method, the accuracy of the centroid in quantifying movements was sacrificed meaning that a direct measurement of movements from the listmode data is not available. However, this is not regarded as too much of a problem as the centroid was designed as a tool for movement detection to enable data reframing and not to quantify movements which would also have required accurate measurement of rotational movements that the centroid does not provide. It was expected that the centroid would also experience movements due to tracer redistribution that would make a proportional relationship between movements of centroid as measured from the listmode versus those measured from co-registration parameters during these periods difficult. The benefits of using the centroid analysis for motion detection was expected to be, and is, from a practical point of view. Post data acquisition, someone familiar with the behaviour of a tracer, that has good binding over the whole of the brain and good coupling between grey matter and white matter regions, could use the centroid as a means of locating where movements are likely to have occurred. Along that line of thought comes the obvious fact that the centroid will perform according to the experience of the user. Different users might come up with different new frame definitions depending on how critical they are about analysing the centroid data. This is a potential issue that needs evaluation.

Recently, (Wardak et al. 2010) have looked into segmenting the head-rest off the transmission image prior to applying AS_FBF. In our work, this was not done for 2 reasons:

Firstly, it was believed that the impact of the head-rest on our scanner bed is not significant on the µ-maps. The head-rest is made up of a thin layer of carbon fibre of average thickness of 2 to 3 mm (measured from low resolution Transmission image as shown in Figure 4.12) and it is not expected that the photons undergo significant attenuation or scatter due to this thin material.
Figure 4.12: Coronal and sagittal image of the head-rest as observed through a low resolution transmission image. Measured thicknesses over different regions (ignoring the blurred regions) show that thickness of the head-rest is between 2 and 3mm, and therefore of importance modest enough to overlook its correction.

Secondly, since all frame-by-frame realignment methods are being compared with the head-rest included, the error due to the latter is expected to be a systematic one in all methods. We should still be able to see the benefits on the improved method based on data reframing. Segmenting the head-rest can be regarded as a potential enhancement to the proposed algorithm that could be looked into in the future.

4.5.3 Current use of the method and prospects

The method has been implemented for a group of early AD patients in an [18F]-FDG-PET study at the WMIC. Its use is currently being considered for extension to amyloid imaging using [18F]-AV-45 as well as [11C]-Diprenorphine in opiate receptor studies. The target group for use of this approach are patients that cannot tolerate wearing a neoprene cap as a tool detection of motion using external instrumentation like the POLARIS, or for scans that have already been acquired but without any measurements for motion. A retrospective reconstruction with motion correction using the centroid analysis is then possible as the centroid is derived straight from the listmode data itself.

Calculation of the centroid in the HRRT is easier as it is octagonal in shape and the listmode data stores the location of the coincident events such that calculation of the centroid is straightforward. However, extending the method for use in a cylindrical
scanner like the Siemens Biograph Truepoint PET-CT, which is also housed in WMIC, is likely to be slightly more complicated as opposing arcs would have to be considered instead of opposing panels. Also, it is known that the scanner stores the location of the LOR associated for each detected event and not the location of the event itself. The calculations are slightly more complicated but since it is beyond the scope of this work on the HRRT (and also currently the work of another person), it will not be discussed any further.

4.6 Conclusions

Frame-by-frame realignment methods are still popular in daily clinical use as they are computationally less intensive than event-based methods and do not require access to proprietary software to be implemented. Moreover, the latter requires use of precise and accurate instrumentation methods for measuring those movements and finding such a system is still work-in-progress. One of the main drawbacks of motion correction methods in literature is that it fails to address the problem of correcting for motion within frames. With listmode data acquisition and storage becoming more and more popular, we have successfully exploited its temporal flexibility to calculate the centroid of activity during a scan and refer to those centroid movements to reframe the data in order to produce frames that contain minimum motion. We showed via qualitative as well as quantitative assessments how the proposed method RAS_FBF performs better than frame-by-frame realignment methods that are conventionally used in clinical environments (FBF and AS_FBF) and suggest the use of the method for studies where external measurements of motion is not available but where there is good binding of the tracer within different regions of the brain and good coupling between grey matter and white matter within the brain.
Chapter 5
Review of Partial Volume Correction Methods in PET

5.0 Introduction

From scanners’ crystals to faster processors, via improved instrumentation, the last 3 decades has seen massive improvements in the development of Emission Tomography. Yet, functional imaging still lags behind structural-imaging-oriented devices like MRI and CT from a spatial resolution point of view due to a number of difficulties like positron range, depth of interaction, non-co-linearity of gamma rays effects. The HRRT, at the time of its introduction, offered the highest possible resolution in brain PET at 2.5mm FWHM (Wienhard et al. 2002), with substantial improvement based on resolution modelling (Reader et al. 2006) bringing this number down to 1.7mm FWHM. A high resolution T1 image, frequently acquired in conventional imaging protocols in clinical setting, has an isotropic resolution of 1mm. More specific regions like the medial temporal lobe, as opposed to a whole brain volume, can be imaged to a higher resolution but requires more scanning time. One of the important aspects of quantitative imaging is to be able to detect subtle physiological changes occurring in the region being investigated. While the major efforts towards improving on attenuation and scatter correction and normalisation are fully justified, there remains the undeniable need to correct for partial volume effects that arise due to the limited resolution of the scanner to obtain an absolute quantification of the feature being imaged.

Partial Volume Effects (PVE) becomes particularly important when the cross-sectional distance of a structure is less than 2 times the FWHM of the scanner’s Point Spread Function (PSF). Several simulation and phantom studies have shown how the regional estimation of radioactive concentration can result in a large bias due to PVE. From this project’s point of view, we are studying regions of the brain that are thin like the
hippocampus in a dementia context. With progressing atrophy, we need to know the true loss of activity within the region of interest that would represent loss in tissue function as opposed to loss in activity due to PVE. The impact of PVE and its correction has been widely reported in literature, especially in conditions where the pathology is accompanied by structural changes within the brain. (Cidis Meltzer et a. 2000; Law et al. 2000; Rousset et al. 2000)

The aims of this review chapter are to understand the Partial Volume problem, especially from a mathematical point of view, and be able to characterise the performance of existing methods in Partial Volume Correction (PVC) that treat it as a series of linear equations. We consider the limitations of methods presented in literature that justify why improvements in existing methods is desirable.

### 5.1 Partial Volume Effect: History, Terming and Characterising

In the late 1970s, (Brooks and Di Chiro 1977) first introduced the concept of ‘Partial Volume Effect’ in X-Ray CT Imaging where crystal size, at over 2cm, produced images where for each resolved unit of the image, only part of the region of interest with a different attenuation coefficient would be present and consequently, the image intensity is partly altered. The problem was further extended to PET and SPECT which have limited resolution in all dimensions as opposed to CT that has very high in-plane resolution but limited axial resolution. Hence, the PVE concept extended to the sampling of small structures within these modalities.

Over the last 2 decades, with improvements in physics and electronics, the term ‘Partial Volume Effect’ has been mostly employed to describe the degree of blurring within an emission tomographic image by a point spread function (PSF) which is the measured distribution of activity resulting from activity at a single point. Physicists and image analysts in the field have often expressed disagreement over the term used for this blurring. Skretting, for example, thinks that it should be termed ‘intensity diffusion’ (Skretting 2009). His logic stems from the fact that if a certain concentration of a certain substance is placed in a 3D space, the molecules of that substance will diffuse following a Gaussian 3D pattern for the concentration of this molecule to spread out. Resolution-
Recovery might be a better term to explain the phenomenon of PVE where the aim is to recover the radioactivity concentrations based on a resolution model for the scanner.

However, it is worth noting that the partial volume problem makes sense when referring to a region (although in the limit, an image voxel can be considered a region) and that within that region, only part of the activity is measured as the signal is lost due to the limited resolution of the scanner. As mentioned earlier, the problem specifically applies to regions whose cross-sectional distance is less than twice the FWHM of the scanner’s PSF. With the scanner’s PSF, as shown in Figure 5.1, not being a step response but rather Gaussian in shape, the activity from the tail of the latter does not form part of the sampling space of the volume of interest, such that the activity region is underestimated. This underestimation of signal persists until the size of the volume of interest is not at least twice the FWHM of the scanner’s PSF in any of its cross-sectional directions.

![Figure 5.1](image)

**Figure 5.1:** Response of the scanner to true concentration of activity (solid line) is a Gaussian (dotted line). Consequently the measured signal has its tails outside the sampling region such that the activity within that region is underestimated as shown by the arrows.

Moreover, if the volume of interest (VOI) is surrounded by, or is in close proximity of, another activity distribution, the true activity measurements within the VOI is likely to be contaminated in what is termed *spillovers* between activity regions. 3 cases of spillovers are possible with two outlined in Figure 5.2 where: (a) the VOI is surrounded by a
background region that is hotter, such that activity spilling from the background into the VOI is higher than the activity lost by the VOI to the background, leading to overestimation of activity concentration within the VOI. (b), the VOI and the background has same activity concentrations such that the spillovers between VOI and background cancel each other. (c), the VOI is hotter than the background such that activity spilling from the VOI into the background is more than activity from background spilling into the VOI, leading to an overall underestimation of the true concentration within the VOI.

Physically, these phenomena manifest themselves as blurring on the PET images, with difficulty in separating 2 small neighbouring objects. Partial Volume Correction (PVC) is therefore the method used to correct for the loss in signal in objects whose sizes challenge the resolution of the scanner, and along that same stride, correct for contaminations in signal measurement among neighbouring regions.

**Figure 5.2:** Illustration of 2 possible cases of spillover effects in a middle region of interest surrounded by (i) a hot background and (ii) a cold background. The true and measured image are both illustrated with profiles drawn across the ROI. For the hot background, blurred profile of the ROI only is given by the dashed black lines; true profile across the top image given by the solid red line and the measured intensity given by the dashed profile, showing how the ROI intensity is overestimated. For the cold background, again black dashed profile shows the mean intensity measured for the ROI in absence of a background, solid blue profile shows the true intensity across image while dashed blue profile shows how the true measurement is underestimated as the activity spilling out of the ROI is higher than the activity from the cold background spilling into the ROI.
5.2 Derivation of the object-image relationship

Consider an ideal linear system, Equation 5.1, where the observed image $g(r)$ is the result of a Fredholm integral equation of the first kind of the unblurred true activity concentration $f(r)$ with the scanner’s spatially variant PSF $a(r, r')$, $r$ and $r'$ being the vectors in image space and object space respectively. $\mathcal{R}(r)$ is a local region around $r$ with significant PSF values.

$$g(r) = \int_{\mathcal{R}(r)} a(r, r') f(r') dr'$$  
 Equation 5.1

A spatially variant PSF implies that the response for a point source of activity is not the same across the Field Of View (FOV) of the scanner. In case of the contrary, then the PSF is referred to as spatially invariant. In this project, the PSF is considered to be spatially invariant across the FOV within a linear system. Hence, $a$ only depends on the difference in coordinates $(r-r')$, leading to equation 5.2 where $g(r)$ is now the result of a convolution between $f(r)$ and $a(r-r')$.

$$g(r) = \int_{\mathcal{R}(r)} a(r-r') f(r') dr'$$  
 Equation 5.2

5.3 The ill-posedness in the deconvolution process for image recovery

The convolution integral from equation 5.2 can also be written in its simplified form

$$g(r) = a(r) \ast f(r)$$  
 Equation 5.3

where * represents the convolution operator.
In Fourier space, the spatial coordinates of $r$ may be represented in terms of their spatial frequencies $u$ such that the relationship between a function $f(r)$ and its Fourier Transform $F(u)$ is given by equation 5.4.

$$F(u) = \int f(r)e^{-2\pi iur} dr$$

Equation 5.4

Similarly, the spatial frequency components $u$ can be used to revert back to the spatial coordinates of the object through equation 5.5.

$$F(r) = \int F(u)e^{-2\pi iur} du$$

Equation 5.5

The convolution theorem states that the Fourier coefficients of a convolution is the product of the Fourier coefficients of each function in the convolution. The Fourier Transform $A$ of the PSF $a$ therefore represents the fractional components of spatial frequencies $u$ that is transferred from the object distribution into the image distribution, and is aptly named the Modulation Transfer Function (MTF). So, the frequency components are modulated by the Fourier coefficients of the PSF.

It is desired that $a$ be a Dirac function such that its Fourier Transform $A = 1$. However, the scanner’s PSF has been shown to approximate a Gaussian with a central peak and elongated tails in emission tomography. For the HRRT, previous work has modelled the PSF as a sum of 2 Gaussians (Comtat et al. 2008). As we go to higher frequencies, undersampling the tail of the PSF implies loss of spatial information in the high frequency area.

It would be ideal to recover this spatial information via a straight deconvolution, which would correspond to a division in the frequency domain (Equation 5.6).
\[ F(u) = \frac{G(u)}{A(u)} \]

Equation 5.6

However, this is not the case as PET images are dominated by noise at higher frequencies, and since \( A \) rolls off faster than the noise in \( F(u) \) it accentuates the noise in the system.

In fact, equation 5.2 is missing an important component which is the noise measurements in the system. The noise originates from the inherent Poisson nature of radioactive decay, and also noise induced in the system due to attenuation, scatter and normalisation. A more realistic representation of equation 5.2 is therefore one that includes noise measurements, as shown in equation 5.7 below.

\[
g(r) = \int_{\mathcal{R}(r)} a(r - r')f(r')dr' + \eta(r)
\]

Equation 5.7

Equation 5.7 presents an ill-posed problem as a small change in the image leads to a large error in the object solution. The challenge of partial volume correction is to restore the object by finding the solution to equation 5.7 while suppressing the noise in the image.

### 5.4 Partial Volume Correction Strategies in PET

Partial Volume Correction methods have been around since the late 1970s / early 1980s. They can be categorised into the following:

1. Computation of Recovery coefficients (RC) for objects of known geometry based on physical phantom experiments. The RC of an object is the ratio of the measured activity to the true activity within that sampling region.
Reconstruction-based methods where the PSF of the scanner is modelled within the system matrix of the iterative algorithm used for PET data reconstruction. It is important to note that while this concept is resolution recovery, it is connected to the partial volume problem which applies to regions within the PET data and recovering the mean activity within that region than improving the resolution of the data.

Methods that are applied post-reconstruction. Most of them require accurate delineation of the region via the complementary use of higher resolution anatomical data like an MRI or CT image. Post-reconstruction methods (therefore applies to images) in themselves can be divided into approaches that occur at a voxel-level or at a regional level based on homogeneity assumptions made across the object and image.

In our works, focus is laid on works that relate to defining homogeneous regions in object space and image space. Since this will form the core of discussions, it will be discussed last amongst the methods available in literature. Prior to that, a brief introduction to the other methods found in literature is given.

5.4.1 PVC with Physical Phantom Experiments

The works by (Hoffman et al. 1979) were the first reported PVC methods applied in PET. The method involved calculating the recovery coefficients from objects whose geometries, as well as position on the PET image, were known. From this work, and the same research group, emerged the prediction of the RCs of several brain structures by approximating them from a series of non-overlapping spheres (Mazziotta et al. 1981). In (Hoffman et al. 1979), the RC was defined as the ratio of observed activity to the true value in the absence of any background activity or activity in neighbouring regions. However, it soon became obvious that, in the presence of neighbouring regions of activity, the spill-out from a hot region should be accounted for as much as the spill-in from a neighbouring colder region into that hot region. This forms the basis of Kessler’s formulation that defined the measured activity in the target region to be the sum of the true activity of target region multiplied by the HSRC (Hot Spot Recovery Coefficient – the RC of the hot spot in an isolated region) plus the multiple of the CSRC (Cold Spot Recovery
Coefficient) coming from the background activity or neighbouring activity (Kessler et al. 1984). Mathematically this can be represented by equation 5.8

\[ g_t = (HSRC \times f_t) + (CSRC \times f_b) \]  
Equation 5.8

Where \( HSRC + CSRC = 1 \),

- \( g_t \) is the measured activity within region \( t \)
- \( f_t \) is the true activity within region \( t \)
- \( f_b \) is the background activity / neighbouring activity

This approach was however likely to work if the neighbouring/background object was fairly large and if it did not warrant justification for PVC. The limitation of this approach also extended to when considering organs of higher anatomical complexities like lesions or atrophied organs whose shape could not be modelled from physical phantoms experiments.

### 5.4.2 Reconstruction-based methods

One of the advantages that iterative techniques like the EMML/OSEM algorithm offers, in contrast to analytical methods like Filtered Back-Projection (FBP), is the ability to include a more complex model of the PET acquisition process through the system matrix. One approach is to include the image blurring within the system matrix, modelled using the PSF of the scanner and which can be used to recover images with higher spatial resolution (Brix et al. 1997; Mourik et al. 2008; Sureau et al. 2008). PSF modelling may be applied in both, projection space (Liang 1994) as well as image space (Reader et al. 2003; Alessio and Kinahan 2006). However, the application of PSF modelling in EMML/OSEM algorithms has brought forward issues that are both, operational and scientific, with one arguably linked to the other.

On the operational front, implementation of reconstruction algorithms with PSF-modelling in a clinical setting is difficult as it requires access to the software that drives
the reconstruction process. The latter is often very limited, depending on contractual agreements with the scanner’s manufacturer.

On the scientific side, as opposed to expectations, these slow converging algorithms have been observed to result in locally smoother images and, more importantly, in Gibb’s artefacts seen as overshoots in the measurement at the borders of regions. Probing into the causes and potential solutions for resolution recovery using PSF modelling is exciting but unfortunately, beyond the scope of this work. Readers are directed to this very good reference (Thielemans et al. 2010) as starting point and into the references there-in for further and deeper understanding of resolution recovery using PSF modelling.

For the above reasons, PVC post-reconstruction is still widely popular as it is applied (as the word suggests) post-reconstruction to images that have been obtained from any scanner and thus, does not require modification of any proprietary software. Resolution modelling in reconstruction is also very difficult to apply retrospectively, especially to image based repositories like Alzheimer’s Disease Neuroimaging Initiative (ADNI) that would require storage of the raw data or local archives within PET centres that have historical formats different to the requirements of current reconstruction algorithms.

### 5.4.3 Post-reconstruction methods for PVC

In practice, the continuous form of the image represented by equation 5.7 does not exist as the observed image \( g(r) \) is sampled as discrete image voxels of intensity \( g_j \). Likewise, the true image \( f(r) \) can be represented by discrete image voxels \( f_i \). Furthermore, by making homogeneity assumptions across both \( f_i \) and \( g_j \), voxels can be grouped together to form homogeneous areas that are represented as sets of areas, namely \( D_i^{f} \) and \( D_j^{g} \) within the object and image respectively. Equation 5.7 can therefore be written as:
Equation 5.9

\[
    g_j = \sum_{i=1}^{N_j} \frac{f_i}{\int_{D_j^g} dr} \int_{D_j^g} \int_{D_j^f} a(r - r') dr'dr + \frac{\eta_j}{\int_{D_j^g} dr}
\]

\[
    \eta_i = \int_{D_j^g} \eta(r) dr
\]

Equation 5.9 presents the problem as a linear one, \( g = Af \) that can be solved using linear algebra depending on the choice of \( D_j^f \) and \( D_j^g \). However, early methods for PVC post-reconstruction did not explicitly formulate the problem in the generalised form \( g = Af \).

Anatomically guided pixel-by-pixel approaches limited themselves to reducing the noise propagation due to the deconvolution process by delineating regions from higher resolution images like MRI or CT and recovering the activity within those a-priori defined regions via deconvolution. A brief overview of these methods is given below.

5.4.3.1 Anatomically guided voxel-by-voxel (or pixel-by-pixel) approaches

The smaller the number of regions assumed within the brain, the easier it is to solve equation 5.9. Predictably, back in the days when scanners’ resolution were 18mm at FWHM, the first attempt to solve equation 5.9 was to assume only 1 region within the brain that carried activity, and none elsewhere (Videen et al. 1988). PVC was carried out on a pixel-by-pixel basis. A binary mask of the whole brain was taken and smoothed to match the resolution of the PET scanner. The binary mask is also applied to the observed image and the corrected image is the ratio of this masked image to the smoothed binary mask. The authors assessed the performance of their method by drawing ROIs that included pure GM surrounded by WM and vice versa, 2 small symmetric regions containing an equal amount of GM and WM (in terms of number of pixels), as well as 2 large symmetric regions. They observed that those regions that had an equal amount of GM and WM pixels recovered to the same activity value. Regions of pure GM surrounded by WM and vice versa did not show good recovery as the mask did not account for heterogeneity within cerebral tissues and regarded both GM and WM as one (the only) cerebral tissue in the brain.
The method was improved by (Muller-Gartner et al. 1992) by including the contribution of WM to PVE. A WM binary mask is first obtained from MR segmentation and the average activity within the centrum semiovale, a region assumed to be essentially WM, measured. The latter is then multiplied by the binary mask and the result blurred with the PSF (Gaussian of 14.5mm FWHM) to create a virtual WM PET image. This virtual WM PET image is then subtracted from the observed PET data to produce what is assumed a real GM PET image. From then on, the method follows that of (Videen et al. 1988), whereby the real GM PET image is divided by the smoothed GM binary mask to produce the partial-volume corrected GM PET image.

(Meltzer et al. 1996) further extended this approach by assuming heterogeneous GM, in an attempt to find the GM activity within a specific VOI. The idea here is to effectively solve equation 5.9 with 4 unknowns, assuming CSF=0, calculating WM as per Muller-Gartner et al. (1992), along with a value for cortical gray matter. This value is assumed representative of the surrounding GM that contaminates the VOI. In their works, Meltzer et al (1996) corrected for GM activity in the amygdala using the equation:

$$T_{amygdala} = \frac{g(r) - T_1 f_1(r) \otimes a(r) - T_2 f_2(r) \otimes a(r) - T_3 f_3(r) \otimes a(r)}{f_{amygdala}(r) \otimes a(r)}$$

Where

- $T_{amygdala}$ is the true activity concentration in the amygdala
- $T_i$ = True concentrations of activity in the defined unknown regions GM, WM and CSF respectively. (with $i = 1, 2, 3$)
- $f_i$ = mask for the respective tissue types
- $a(r)$ is the PSF of the system

However, calculating the true activity within a VOI becomes more and more difficult as heterogeneity within the neighbouring region increases.
5.4.3.2 Correction by assessing signal degradation over the mean regional concentration of regions.

An alternate approach to the Muller-Gartner method is to evaluate the impact of the loss in signal due to the limited resolution of the scanner on the mean concentration of activity within homogeneous regions defined in both object space $D_i^f$, and image space $D_j^g$. This forms the basis of the methods of (Labbé et al. 1996) and Rousset et al. 1998), who, based on their choice of $D_i^f$ and $D_j^g$, present the generalised form of equation 5.9, $g = Af$, under a linear framework and find the solution based on the presented framework.

In (Rousset et al. 1998), the homogeneous regions defined in the true image space are the same as those defined in the measured image space. Hence, if $g_j$ is the mean activity distribution in the observed image, and $D_i^f = D_j^g$, $A_{ij}$ takes the form of a square matrix that can be readily inverted such that $f_i$, the mean activity concentration in the object is given by inverting the square matrix $A_{ij}$

$$f_i = (A_{ij})^{-1}g_j$$

Equation 5.10

Where

$$A_{ij} = \frac{1}{VOX_j} \int_{D_j^g} \int_{D_i^f} a(r - r') dr' dr$$

And $VOX_j$ = the number of voxels in the region defined by $D_j^g$.

In the works of (Labbé et al. 1996) on the other hand, the homogeneous regions are defined only in the object space $D_i^f$ with the image space $D_j^g$ sampled as voxels. This results in matrix $A_{ij}$ where $j > i$, hence an over-determined system that can be solved under an Ordinary Least Squares framework, as shown in equation 5.11.
\[ f = (A^T A)^{-1} A^T g \]  

**equation 5.11**

In the limit that both \( D_i^f \) and \( D_j^p \) are sampled as voxels, the problem becomes a straight deconvolution, which as discussed earlier, becomes ill-posed due to noise propagation. However, the deconvolution process becomes more feasible under an iterative framework like that of Van-Cittert (VC) Iterative Deconvolution, given by **equation 5.12** or Lucy Richardson (LR) Iterative Deconvolution, presented under an EM framework in **equation 5.13**.

\[ f^{k+1} = f^k + \alpha(g - Af^k) \]  

**equation 5.12**

Where \( \alpha \) is a regularisation term that is often taken to be 1

\( f^k \) is a vector of the true activity concentration being estimated where

\( k \) is the superscript denoting the \( k^{\text{th}} \) iteration

\( g \) is a vector of the observed image

\[ f^{k+1} = \frac{f^k}{A 1_{N_j}} A^T \frac{g}{Af^k} \]  

**equation 5.13**

Where \( 1_{N_j} \) is a column vector of ones.

The main difference between VC’s method and LR’s method is that the update in the former is additive while it is multiplicative in nature in LR. The impact of the nature of the updates is discussed further below.

(Teo et al. 2007) looked into using Van-Cittert’s algorithm for PVC of the uptake of \([^{18}\text{F}]-\text{FDG}\) and \([^{11}\text{C}]-\text{CO}\) in small tumours. The authors observed good recovery in their phantom
studies on spheres filled with [18F]-FDG with 100% and 103% recovery in spheres that were between 13 mm and 29 mm in diameter and 87% recovery in the smaller spheres (8 mm). However, as highlighted by (Boussion et al. 2009), VC also introduced high levels of noise that explains why (Teo et al. 2007) were constrained to propose the use of their algorithm on tumours that were a-priori well delineated from a higher resolution image. (Tohka et al. 2008) on the other hand, proposed a modified VC algorithm that included reblurred VC method together with an added total variation-based regularisation. Since their Monte Carlo simulated PET images were reconstructed using FBP, they also had to use a Hann filter to suppress the noise amplification that occurred during the deconvolution of the images. (Boussion et al. 2009) proposed a wavelet algorithm that denoises the residuals prior to the update taking place. They assessed the performance of both VC and LR with and without wavelet denoising and essentially observed that the deconvolution step in VC’s method amplified the noise by 3 times more than LR. Application of the wavelet denoising filter did not suppress the high noise amplifications. Resulting Partial Volume Corrected images were difficult to interpret with artefacts around high-intensity regions as well as varying intensities within uniform regions. They concluded that LR was a better method for PVC using iterative deconvolution as it produced images that were qualitatively and quantitatively better, without altering the dedicated voxel intensities within regions, irrespective of their sizes. The reduced noise amplification in the LR approach can be acquainted to the fact that the update in the deconvolution process is multiplicative in nature. Hence, it is more robust to noise propagation than in VC where it is additive. Furthermore, LR is based on the assumption that each activity concentration within every voxel of the image comes from a Poisson distribution. A PET image itself, stems from radioactive decay which is a random process with a Poisson Distribution and reconstruction algorithms like MLEM-OSEM are based on optimising a cost function that is essential based on Poisson Maximum Likelihood. As a first order approximation, the variance is proportional to the mean for MLEM type algorithms which is definitely not the case for FBP images (Barrett et al. 1994).
5.4.3.3 Energy Multiresolution Analysis

Another post-reconstruction method that is gaining more and more interest from research groups are methods using Mutual Multiresolution Approaches (MMA) where a higher resolution PET image is obtained by decomposing an anatomical co-registered image from MRI or CT in terms of their frequencies (high and low frequencies) using Wavelet Transform (WT) and including the higher frequency information (edges) into the PET image. However, wavelet coefficients from a CT or MRI image do not have the same levels of spatial and intensity details as the corresponding PET image, such that their inclusions into the PET data is not straightforward. Key papers in literature that tackle this problem are by (Boussion et al. 2006), where high frequency information from a co-registered CT or MRI image is extracted using WT and scaling the higher frequency details to compensate for the difference in spatial details that both imaging modalities offer; and (Shidahara et al. 2008) where on top of accounting for the difference in spatial resolution, the authors also accounted for an intensity scaling factor between the anatomical and functional image as well as a weighing factor between anatomical information versus the noise in the system. Work in this field is still in progress as to whether the relationship in spatial details between anatomical and functional images are global and linear or not and how to gauge the difference between high frequency anatomical details and noise. Since MMA methods are beyond the scope of the present work, it will not be discussed any further and more information can be obtained from the above mentioned papers and the references there-in.

5.5 Critical Appreciation of post-reconstruction PVC methods

Methods of (Rousset et al. 1998 and Labbé et al. 1996) are essentially region-based methods where, as the name suggests, regions have been defined from an anatomical image based on homogeneity assumptions. In fact, the widely acknowledged method of (Rousset et al. 1998) is based on the critical assumption that the segmentation of regions within the image volume is accurate and complete; if not within the whole volume, at least within the sub-volume of the image under investigation. This means that ignoring other sources of error like registration, the regions identified should contain no variability within them. In our study of AD patients where the focus is on quantifying regions of the
Medial Temporal Lobe: a region that is key to changes, both in terms of metabolism and atrophy. One automatic way of delineating regions over the Medial Temporal Lobe is by normalising to an Atlas like the Hammers’ Atlas (Hammers et al. 2003) to the co-registered MRI image. It is desired to quantify the metabolic changes in the hippocampus. Inspection of the Hammer’s Atlas shows that it is surrounded by 7 other regions. However, the regions specified within the Hammers’ Atlas is not complete as the hippocampus is also surrounded by white matter tracts, CSF from atrophied tissues as well as from the lateral ventricles. The latter also houses the choroid plexus which carries a lot of blood vessels. All these contribute to the heterogeneity of the Medial Temporal Lobe and separating them into smaller and smaller regions is difficult, if not impossible, hence leading to inaccuracies in the recovered signal. Increasing the number of defined regions surrounding the hippocampus would help in improving the accuracy of the measured signal. However, with increasing number of regions, the A matrix becomes larger and larger such that its inversion becomes an ill-posed problem that introduces a bias into the region of interest.

In the limit that all voxels are defined as individual regions within the image, hence making no assumption of homogeneity within, the problem is a deconvolution one that can be solved using VC or LR. However, as already discussed above, VC is not recommended for PVC. The slow-converging LR algorithm under an EM framework, on the other hand, has shown to be prone to Gibbs artefacts at the borders of homogeneous regions.

Last, but not the least, it is required that PVC methods developed in emission tomography account for the noise properties of the image. As mentioned earlier, radioactive decay is a random process that follows a Poisson distribution. Other forms of noise get introduced into the system. For example, systematic errors such as errors in the system matrix result in bias with the images and changes the propagation of noise. (Aston et al. 2002) essentially simplifies the noise structure by separating them into 2 models:

Firstly, uncorrelated noise models where, as the name suggests, there is no spatial correlation in the noise distribution across the image such that the noise is effectively random additive Gaussian noise across the image. The authors of this paper detail out
how the methods of (Labbé et al. 1996) and (Rousset et al. 1998) are implicitly based on this model and show how, for either methods, despite producing good estimates of the regional concentration of activity, but underestimates or overestimates the standard deviation associated with each estimate. The authors also highlight what has been discussed earlier as that both Labbé’s and Rousset’s method are essentially addressing the same problem except that different weights are applied to the least squared problem to obtain the desired solution.

Secondly, correlated noise models where one of the components of the noise model is due to the scanner’s PSF where the signal itself, or the signal combined with other noisy processes, are effectively blurred by the PSF.

The authors considered each of the above noise models individually and a combination of them to show that a more realistic error associated to the recovered estimate is obtained when a combination of both noise models is used.

One of the other key aspects explored by the authors is the testing of the validity of homogeneity assumptions within the defined regions. They showed that in practice, the task proved to be too difficult due to a lack of computational power as the defined regions increased in size and numbers as well as to being unable to formulate an accurate noise model for the system. They also reflect on the fact that while they were successful in showing the importance of correctly modelling the noise in their simulation studies, their measured datasets did not show the same results in terms of retrieving the correct associated error for the simple reason that it was difficult to accurately characterise the noise model in the PET image.

5.6 The Partial Volume Effect and Correction in perspective

This chapter has so far given a concise description of the problem of PVE in PET and has outlined several methods used to address the problem. Table 5.1 gives a summary of methods that have been described so far. It does not include the multi-resolution approach via wavelet decomposition as it cannot be explicitly derived from Equation 5.9.
These are the core methods that have been derived for use in a clinical or research setting, based on assumptions like homogeneity of regions within the brain, or properties of the emission image (e.g., activity following a Poisson distribution, preservation of the Poisson characteristics with MLEM-OSEM reconstruction, noise modelling within the image, whether implicit or explicit).

<table>
<thead>
<tr>
<th>Method / Authors</th>
<th>$D^f_i$</th>
<th>$D^g_j$</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rousset et al. (1998)</td>
<td>ROIs</td>
<td>ROIs</td>
<td>$f = A^{-1}g$</td>
</tr>
<tr>
<td></td>
<td>(N)</td>
<td>(N)</td>
<td></td>
</tr>
<tr>
<td>Labbe et al. (1996)</td>
<td>ROIs</td>
<td>Voxels</td>
<td>$f = (A^TA)^{-1}A^Tg$</td>
</tr>
<tr>
<td>Lucy Richardson</td>
<td>Voxels</td>
<td>Voxels</td>
<td>$f^{k+1} = \frac{f^k}{A1+n_j}A^Tg Af^k$</td>
</tr>
<tr>
<td>Van Cittert’s</td>
<td>Voxels</td>
<td>Voxels</td>
<td>$f^{k+1} = f^k + \alpha(g - Af^k)$</td>
</tr>
</tbody>
</table>

Table 5.1: Summary of post-reconstruction PVC methods in literature.

Over the past decade, PVC from a post-reconstruction point of view has been based on assessing the limitations of the above methods and proposing variants of the above algorithms in an attempt to refine the algorithm and obtain a more accurate solution. (Quarantelli et al. 2004) implemented a software that provided PVC using 4 methods, the method by (Videen et al. 1988), the method by (Muller-Gartner et al. 1992), the one by (Rousset et al. 1998) and the fourth one, a variant of the Muller-Gartner method, where the white matter tissue concentration is calculated using Rousset’s method instead of being estimated from a large region of white matter assumed to be devoid of PVEs (Rousset et al. 1998). The software allowed the authors to inter-compare the methods in terms of their accuracy (via Grey Matter Recovery Coefficient) and precision (Via Grey Matter Coefficient of Variation) over varying magnitudes of errors in segmentation, co-registration and resolution estimates. Those were simulated on 4 virtual PET phantoms derived from their MR images and generated to simulate the properties of the ECAT
EXACT 47 Siemens PET scanner. They found that Rousset’s method out-performed the 3 other voxel-based methods both in terms of accuracy and precision of the true concentration of radioactivity. The modified Muller-Gartner method was the most accurate of the 3 voxel-based methods. They also observed that errors in co-registration affected the accuracy of the methods the most as they produced the highest coefficients of variation in the measurements. While (Quarantelli et al. 2004) highlight image co-registration as the main problem in achieving a good signal recovery, (Frouin et al. 2002) pointed at image segmentation as being more significant. The authors showed that if the volume of the Caudate Nucleus is missed by 25%, its Apparent Recovery Coefficient (ARC) would decrease by 5%. The segmentation problem is effectively linked to defining homogeneous regions within the brain; something which is difficult for reasons already mentioned above.

In terms of variants of existing methods proposed, the modified Muller-Gartner (mMG) method has already been looked at. Most papers tend to propose a modified Rousset’s approach where the A matrix is based on what can be termed as a ‘refined segmentation’. (Svarer et al. 2005) have devised an automatic way of delineating VOIs templates for use in analysis of images from emission tomographs. The templates initially stem from manually drawn VOIs taken from high resolution MR images that are then transferred and warped onto the subjects’ brains for further analysis. These VOIs are those that can be used for creating Rousset’s A matrix for PVC. (Maroy et al. 2009) considers only the signal coming from limited number of voxels for each defined region so that the noise within the system is suppressed. Very recently, (Thomas et al. 2011) produced a variant of the mMG method by performing a voxel-based correction on a PET image in which spillovers between parcellated regions has already been accounted for. En passant, in the same work, the authors also compared their method with VC deconvolution and showed that the latter method produced the worst recovery for GM values.
5.7 Summary of literature review findings:

PVEs occur due to the limited spatial resolution of the PET scanner. Methods for PVC can be divided into several groups. Those groups that perform PVC post-reconstruction include voxel-based deconvolution methods like LR or VC, and those methods that use image-based segmentation like those summarised in Table 5.1. It has been suggested that LR is a better voxel-based deconvolution method than VC. Among the image segmentation-based methods in Table 5.1, Rousset’s method has proved to be very popular in clinical research. However, the method relies on a complete and accurate segmentation of the volume, or at least a sub-volume, of the image.

With segmentation an imprecise science due to the brain’s complex heterogeneous structure, the chances of tending towards the exact solution gets bigger when every voxel surrounding the region of interest is considered an individual region. However, adopting such a philosophy makes the inversion of the A matrix in Rousset’s method an ill-posed problem that will eventually introduce a bias in the partial volume corrected measurement. Also, with higher resolution scanners like the HRRT, a higher precision is required for the segmentation procedure.

In the event that every voxel within the image volume was a region, the problem would have been a deconvolution one where voxel-based methods like the LR iterative deconvolution can be used to address the partial volume problem. This slow converging algorithm under an EM framework works well in an environment where the data is essentially Poisson in nature, as long as the PSF is well characterised. It does not require segmentation of the image volume but is prone to Gibb’s artefacts at the borders of homogeneous regions.

It would be ideal to have an algorithm that amalgamates the advantages of both region-based methods like Rousset’s method and voxel-based methods like the LR iterative deconvolution method. Where the image is segmentable, the algorithm would output a mean concentration of activity for that region following the assumptions of homogeneity as does the method of (Rousset et al. 1998). The rest of the image will still undergo partial
volume correction but since an accurate segmentation is unavailable over that region, it would undergo a voxel-based one.

We have designed and implemented a simple algorithm that combines the characteristics of both region-based and voxel-based methods for partial volume correction. The next chapter explains the mathematical and logical aspect of the algorithm and evaluates the performance of the algorithm against Rousset’s region-based method and the LR iterative deconvolution method under an EM framework, both of which it is a hybrid of.
Chapter 6
Introducing PARSLR:
A hybrid between region-based and voxel-based methods for partial volume correction in emission tomography.

6.1 Introduction

We are going to create an algorithm that works as a hybrid between region-based and voxel-based methods for PVC. We start off with describing the Lucy Richardson (LR) iterative deconvolution method from scratch: the most acknowledged method for purely voxel-based methods found in literature (Boussion et al. 2009; Tohme and Qi 2009; Thomas et al. 2011). We will then modify the LR method to allow the algorithm to consider homogeneity assumptions where possible, in the way region-based methods do (Labbé et al. 1996; Rousset et al. 1998). The end product is a PARtially Segmentable Lucy Richardson (PARSLR) algorithm that allows for segmentation of the region(s) of interest within an image volume or sub-volume where this can be accurately achieved while leaving the remaining voxels within the image volume to undergo conventional LR iterative deconvolution during the partial volume correction process.

We will then assess the performance of PARSLR with respect to the methods of (Rousset et al. 1998) and LR iterative deconvolution to see how well PARSLR performs in conditions where the latter methods are known to experience problems in finding an accurate corrected measurement.

6.2 Problem Definition

Sticking to the nomenclature of Chapter 5, let \( F \) be a 3D PET image volume of dimensions \( N_I, N_J, N_K \) describing the true non-blurred radioactivity distribution. The matrix will be
index by \(i, j, k\) from 1 to \(N_i, N_j, N_k\) respectively. Let \(Q\) be the 3D kernel representing the blurring, which is assumed to have equal dimension in \(x, y\) and \(z\) and will be indexed from \(-N_Q\) to \(N_Q\) in each dimension. Finally, let \(G\) be the resulting 3D PET image volume of dimensions \(N_i, N_j, N_k\) produced by convolving the original image with the kernel. For clarity, this matrix is indexed by \(l, m, n\).

### 6.3 Construction of the problem

\[
G = Q \ast F \quad \text{Equation 6.1}
\]

\[
G_{l,m,n} = \sum_{i=\max(1,l-N_Q)}^{\min(N_i,N_Q+l)} \sum_{j=\max(1,m-N_Q)}^{\min(N_j,N_Q+m)} \sum_{k=\max(1,n-N_Q)}^{\min(N_K,N_Q+n)} Q_{l-i,m-j,n-k} F_{i,j,k}
\]

Vectorising \(F\) and \(G\) turns the problem into a simple linear one;

\[
g = Af \quad \text{Equation 6.2}
\]

\[
A_{l,m,i,j,k} = \begin{cases} Q_{l-i,m-j,n-k} & -N_Q \leq l - i \leq N_Q; -N_Q \leq m - j \leq N_Q; -N_Q \leq n - k \leq N_Q \\ 0 & \text{otherwise} \end{cases}
\]

Under an EM framework, the Lucy-Richardson iterative deconvolution is presented as

\[
f^{p+1} = f^p A^T 1_{N_{ijk}} A^{T^{-1}} g / Af^p 
\]

\text{Equation 6.3}

where,

\(f^p\) is the estimate of \(f\) at the \(p^{th}\) iteration; \(g\) is the measured blurred data, \(1_{N_{ijk}}\) is a column vector of ones and of dimensions \(N_i \times N_j \times N_k\). The division operations in equation 7.3 refer to element-wise divisions. For multiplication processes, if the first term is a matrix, then the process refers to a matrix multiplication. Else, where the first term is a vector, it refers to element-wise multiplication.
6.4 Efficient implementation of the LR algorithm under an EM framework

While the element-wise operations in Equation 6.3 are mathematically trivial, the computation of $Af$ and $A^Tg$ in an optimised way needs to be looked into, as a sparse representation of $A$ would be inefficient due to replication of terms in $A$. By inspection, it is quite obvious that computation of $Af$ is a straightforward convolution operation.

Calculating $A^Tg$:

Let $f' = A^Tg$,  

$$f'_{i,j,k} = \sum_{l=\max(1,i-N_Q)}^{\min(N_I,N_Q+i)} \sum_{m=\max(1,j-N_Q)}^{\min(N_J,N_Q+j)} \sum_{n=\max(1,k-N_Q)}^{\min(N_K,N_Q+k)} Q_{l-i,j-m,k-n} G_{l,m,n}$$

Equation 6.4

From equation 6.4 above, using $A^T$ is equivalent to convolving with the flipped kernel where every dimension of $Q$ is flipped. For the case of a symmetric kernel, the transposed and forward kernel are the same such that the kernel used is still $Q$.

6.5 Reading of the LR Iterative Deconvolution Algorithm

The Lucy-Richardson iterative deconvolution algorithm is read as follows:

1. Choose an initial image $f^p$ ($p=0$);
2. Convolve with the kernel $Af^p$;
3. Ratio with the measured data $\frac{g}{A^p f}$;
4. Convolve the result with the flipped kernel. For symmetric kernels, it is identical to the original blurring operation. Let $r = A^T \frac{g}{A^p f}$;
5. Update by performing element-wise operation $f^{p+1} = \frac{f^p r}{A^{T^1 N_{IJK}}}$;
6. Repeat from Step 2.
7.6 Introducing PARSLR

We modify the LR algorithm to allow for the use of homogeneous segmented regions where this is can be achieved accurately within the image volume. Where such segmentation is impossible, $D_i^l$ is left as image voxels. Specifically, with the image voxels $D_j^g$, let

$$g = AB\theta$$  \hspace{1cm} \text{Equation 6.5}

Where,

$\theta$ is a vector of concentrations within regions and

$B$ is a matrix of ones and zeros that maps these concentration to image voxels, with

$$B_{i,v} = \begin{cases} 1 & \text{if the voxel } i \text{ is within region } v \\ 0 & \text{otherwise} \end{cases}$$

For example, Figure 6.1 shows a 2D image made up of 4x4 pixels. The image is segmentable into 4 distinct regions shown in red, blue, yellow and green. However, one of the pixels (shown in black) cannot be defined as being part of any of the surrounding regions. In the corresponding $B$ matrix, we therefore have 5 column vectors of 16 rows each, with 4 of them mapping each of the regions in red, blue, yellow and green. Where a pixel cannot be tagged to a region, it is left as a region as shown in black.

From equation 6.3, we are effectively substituting $AB$ for $A$ and $\theta^p$ for $f^p$. The resulting equation for the modified LR algorithm is thus:

$$\theta^{p+1} = \frac{\theta^p}{B^TA^T1_{NJK}}B^TA^Tg/AB\theta^p$$  \hspace{1cm} \text{Equation 6.6}

The above algorithm has been designed to accommodate for segmentable homogeneous regions where possible and has been branded (PARtially Segmentable Lucy Richardson).
Figure 6.1: 2D 4x4 pixels image (right hand side) and its corresponding B matrix (left hand side). Image is segmentable into 4 regions (red, blue, yellow and green) and a pixel which cannot be assigned to any region (black). It is therefore left as an individual region. Matrix B is indexed counting rows across columns, in the same way that linear indexing works on MATLAB.

6.7 Reading of PARSLR

Since $A^T 1_{N_{ijk}} = 1_{N_{ijk}}$ will occur for kernels of unit area and for areas well within the image boundaries such that the denominator in step 5 of the LR deconvolution drops out. $B \theta$ equates to populating the image with the defined intensities. The term $B^T$ in the numerator equates to summing up intensities within regions, while the final $B^T 1_{N_j}$ term in the denominator is the number of voxels in a region. The algorithm of equation 6.6 summarises to:

1. Choose an initial concentration within each region or image voxel $\theta^p (p=0)$. This could be the mean concentrations of the homogeneous regions;
2. Map these values to the image using segmented regions defined by the matrix $B$;
3. Convolve resulting image with kernel $AB\theta^p$;
4. Ratio with the measured data $\frac{g}{A B B^p}$;

5. Convolve the result with the flipped kernel $r = A^T \frac{g}{A B B^p}$;

6. When calculating the numerator $B^T r$, the inner product operation effectively leads to summing all the elements of the vector $r$ for every region defined in $B$, while the inner product operation in the denominator $B^T 1_{N_j}$ is summing all the voxels across every region defined in $B$. Hence, we ultimately end up in calculating the mean within every region the image $r$ by performing $w = \frac{B^T r}{B^T 1_{N_{i,j,k}}}$;

7. Update the regional concentrations through multiplication $\theta^{p+1} = \theta^p w$;

8. Repeat from step 2.

### 6.8 Performance Assessment of PARSLR as a method for PVC

#### 6.8.1 Philosophy of assessment

We assessed the performance of PARSLR with respect to the region-based method of (Rousset et al. 1998) and the voxel-based method of LR. We want to test:

1. Whether in noiseless data, and in an environment of complete and accurate segmentation, all 3 algorithms are in agreement with each other.

2. How each algorithm behaves with multiple noisy realisations of the data in an environment of complete and accurate segmentation.

3. In an environment of incomplete segmentation, where we are able to delineate only the region of interest but unable to identify the neighbouring regions accurately. We therefore leave those voxels as either:
   a. Being an individual region (each voxel); as we want to test how the increase in number of regions affects the performance of Rousset’s algorithm and how PARSLR behaves in comparison.
   b. One homogeneous background. One solution to prevent the ill-posed nature caused by too many regions in the A matrix of Rousset’s method is to define the neighbourhood as one homogeneous background. Such a
step would create a bias into the partial volume corrected signal in the event that the background is effectively heterogeneous.

6.8.2 Materials and Metrics used for assessment

We applied the partial volume correction methods on:

6.8.2.1 A Simulated hippocampus phantom

A standard region template (The Hammers’ template) was spatially normalised (using SPMS software) to the T1-weighted MRI image of a healthy elderly volunteer (male, 64 years old) and the 7 regions surrounding the hippocampus as according to the Hammers’ Atlas (Hammers et al. 2003) identified, as shown in Figure 6.2. The corresponding co-registered $^{18}$F-FDG PET scan was then intensity normalised to the hippocampus and the regions scaled up by a factor of 5. Since most regions around the hippocampus have intensities that are comparable to each other, the intensities of some regions were changed on purpose to better simulate for partial volume effects between the hippocampus and the neighbouring regions. The regions and their respective intensities are given in Table 6.1. The phantom was tested under

a. No-noise condition where the phantom image was only blurred by a Gaussian of 2.5mm FWHM that is representative of the HRRT’s PSF. No forms of noise were applied to the data.

b. Multiple levels of noise representative of counting statistics present in a PET image, by adding random Poisson Noise to the phantom (using poissrnd, the standard function available in Matlab). Addition of different amounts of Poisson noise to the data is achieved via first scaling the image with a scale factor, adding Poisson noise and rescaling the result back to the original scaling as explained mathematically in equation 6.7 below.

$$ g = A^{(\frac{r f + \eta}{r})} $$

Equation 6.7

where,

$\eta$ represents the Poisson noise added to the system where
\[ \eta \sim \text{Poisson}\{\lambda \} \]

\( f \) represents the image where the variance \( \lambda \) is equal to the mean of \( f \)

\( \tau \) is the scaling factor

\( g \) is the resulting image after rescaling and smoothing with a Gaussian of 2.5mm at FWHM to represent the scanner’s PSF \( A \).

The values used for \( \tau \) were: 0.1 to 1.0 in increasing increments of 0.1; 10 and 100. For each noise level, 30 realisations of the image were generated. The noise model used is therefore one that includes only pure correlated noise as per (Aston et al. 2002).

![Figure 6.2: Left hippocampus from Hammers’ Atlas and its immediate surroundings used as mathematical phantom. Regions indexed between 1-8 as according to colourbar and detailed in Table 6.1 below](image)

<table>
<thead>
<tr>
<th>Region Index</th>
<th>Region Name</th>
<th>True Value</th>
<th>Value Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hippocampus</td>
<td>5.00</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>Amygdala</td>
<td>4.95</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Parahippocampal and ambient gyri</td>
<td>5.15</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Superior temporal gyrus, posterior part</td>
<td>7.12</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>Fusiform gyrus</td>
<td>5.45</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>Insula</td>
<td>6.39</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>Posterior temporal lobe</td>
<td>6.60</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>Lateral ventricle, temporal horn</td>
<td>3.91</td>
<td>4</td>
</tr>
</tbody>
</table>

| Table 6.1: Regions that surround the hippocampus. Each region has been allocated a value that is representative of the median activity value recorded on FDG-PET scan, intensity normalised to the hippocampus and scaled by a factor of 5. Some regions (e.g. amygdala) have had their values substantially changed to enable examination of the changes due to spillovers between adjacent hot and cold regions. The true values and values used are both indicated in the 3rd and 4th columns. |

142
The realisations were corrected for partial volume under the philosophy described in the previous section, illustrated in Figure 6.3 and summarised in Table 6.2; the latter describing how the resulting graphs have been named for ease of comparison. It can be noted from Figure 6.3b and 6.3c that only a sub-volume of the image is used for the assessment under Rousset’s method in the event that every voxel is considered an individual region. This is because the inversion of Rousset’s A matrix is limited by computational power, on top of the ill-posedness problem of inverting large matrices. Even for this limited number of regions, a Moore’s Penrose pseudo-inverse had to be used for the inversion. However, for PARSLR, the whole image volume is always used as the efficient implementation of the algorithm allows its computation with the same computational power. The metrics used for the assessment are bias, variance and mean squared error (MSE) between realisations at different noise levels compared to the true value within the hippocampus region of the phantom.

<table>
<thead>
<tr>
<th>Cases</th>
<th>Rousset</th>
<th>LR</th>
<th>PARSLR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete and accurate Segmentation</td>
<td>Rousset_complete</td>
<td>LR</td>
<td>PARSLR_complete</td>
</tr>
<tr>
<td>Hippocampus and remaining voxels set as individual regions</td>
<td>Rousset_vox</td>
<td>LR</td>
<td>PARSLR_vox</td>
</tr>
<tr>
<td>Hippocampus and remaining voxels set as homogeneous background</td>
<td>Rousset_back</td>
<td>LR</td>
<td>PARSLR_back</td>
</tr>
</tbody>
</table>

*Table 6.2: Methods of performance assessment for the 3 PVC methods. PARSLR_back and LR back are obviously not included/required because these algorithms undergo a voxel-based deconvolution for PVC outside the segmented region and therefore does not require a region-based one.*
6.8.2.2 Measured Data: 2D brain slice phantom

Measured data for the assessment in the form of a 2D printed brain phantom, designed by Dr. Pawel Markiewicz, Physics research fellow at the WMIC. In PET scans involving patients, the true concentration of activity within a region or its surroundings is not
known such that it is impossible to gauge the accuracy of the method in recovering the true value. However, what can be done is to simulate the partial volume effects in the form of a physical phantom and check whether the corrected activity concentrations are in line with the measured data. We can then infer from these observations to explain the changes observed when applying the PVC methods to patients’ data.

A transaxial slice of an MR image was segmented into Gray Matter (GM), White Matter (WM), scalp and Cerebrospinal Fluid (CSF). This segmented slice was then printed on a sheet of paper with 14.3 MBq of radioactive Fluorine-18 ink in descending order of ink density (CSF=0). This sheet of paper was then scanned on the HRRT for 360 minutes and reconstructed with OP-OSEM, 16 subsets, 10 iterations for the following frame definitions: 12x5s, 12x10s, 10x30s, 10x60s, 5x120s, 6x300s, 1x3600s, 1x14400s, and 1x119s (the final 1x119s frame was not used in further experiments). The data was divided according to this frame definition as we also wanted to observe the effect of counting statistics on PET images that have been corrected for partial volume effects. A plot of the number of trues in each of the above mentioned frames is given in Figure 6.4 below:

![Number of trues in each frame](image)

**Figure 6.4:** Number of trues in each frame for frame1 to frame55. Number of trues for frame56=1635 millions, and frame57 = 2842 millions are not shown as they go higher up the scale. Frame of 119s was discarded and not shown on the graph.
The MR slice and the resulting printed image are shown in Figure 6.5a. From this printed and reconstructed phantom, 2 regions were selected to assess the performance of the PVC algorithms. Region1 shown in the red contours of Figure 6.5a which is effectively a GM region and Region2, shown in red contours in Figure 6.5b is a CSF region. Again the whole image volume could not be used for calculating Rousset_vox as the resulting $A$ matrix is still too large for inversion, even under the use of a Moore’s Penrose pseudo inverse. The sub-volumes used for each region are delineated in green. However, for both regions, the assumed homogeneous background used was smaller, as shown with the yellow contours in their respective figures as we wanted to include regions that would be affected by significant PSF values only. Yet, it is clear from Figure 6.5a and 6.5b that the surroundings of the selected regions is much bigger than the locus of significant PSF values. This was done in a deliberate attempt to increase the heterogeneity of both regions as Region1 is surrounded by essentially WM while Region2 is bound by essentially GM.

At the end of the scan, the activity concentration in the delineated Region1 was measured by punching holes off the sheet of paper and measured in a well-counter. Region1 was observed to have an activity concentration of 1710 K bq/mm$^2$ after correction for background counts and decay. Region2 was known to be zero as it was not printed with radioactive ink. PVC was applied to Region1 and Region2 using the same philosophy summarised in Table 6.2 to see which method gave the smallest error in recovery of that region.
6.9.1 Simulated hippocampus phantom

First, the performance assessment was carried out on the simulated hippocampus phantom under conditions of no noise before exploring how the algorithm behaves under different levels of noise firstly under a complete and accurate segmentation and then when only the hippocampus can be accurately segmented. The results are detailed below.
6.9.1.1 Noiseless data – Complete and accurate segmentation

Under a complete and accurate segmentation and under a noiseless environment, it can be observed, via Figure 6.6, that Rousset, LR and PARSLR all recover the true values of regions within the hippocampal phantom.

Figure 6.6: Top: Images showing image intensities and profile locations; Bottom: Profiles of image intensities. Image a (green profile) = original unsmoothed image; image b (red profile) = image smoothed with 2.5mm Gaussian kernel; image c (blue profile) = recovered image intensities following 100 LR iterations; image d (pink profile) = recovered image intensities following Rousset PVC with a complete and accurate segmentation; image e (black profile) = recovered image intensities following 100 iterations of PARSLR PVC with a complete and accurate segmentation. The green, pink and black profiles overlie each other.
The image from Figure 6.6c shows the recovery of the image following 100 LR iterations. Gibb’s artefacts, observed as overshots at the boundary of regions, are visible at the borders of the regions within the phantom and are clearer on the blue profile at the bottom of the figure. In recovery using PARSLR, those Gibb’s artefacts are suppressed.

6.9.1.2 Multiple levels of correlated Poisson noise

6.9.1.2.1 Complete and accurate segmentation

Figure 6.7 depicts the results of PVC applied with the methods of Rousset and PARSLR under complete and accurate segmentation and compared with LR over different levels of simulated Poisson random noise. It can be observed that both, Rousset and PARSLR recover the mean intensity of the hippocampus (True value = 5, from Table 6.1) to approximately the same level. In fact, Rousset, under a complete and accurate segmentation, seems to perform slightly better than PARSLR with increasing noise levels (at lowest noise level with COV of 4.5%, PARSLR = 4.992±0.011 and Rousset = 4.992±0.018; at noise level with COV = 54.8%, PARSLR = 4.992±0.061 and Rousset = 5.004±0.061; and at highest noise level with COV = 144.2%, PARSLR = 4.957±0.163 and Rousset = 4.981±0.171). LR, for 50 iterations, performs slightly under both Rousset and PARSLR with a mean intensity of 4.86 on average over the noise levels applied. All 3 algorithms, for the simulated case, show preserved noise properties within the image as the variance of the simulated regions across realisations remains roughly the same between uncorrected and corrected data for all methods.
Figure 6.7: Graph of mean intensity of simulated hippocampus versus different Poisson random noise levels ranging from High (COV = 144.2%) to Low (COV = 4.5%) under a complete and accurate segmentation for (i) uncorrected hippocampus (green triangles), (ii) corrected with LR, 50 iterations (LR - black inverted triangles), (iii) corrected with Rousset’s method (Rousset Full – red circles) and (iv) corrected with PARS LR, 50 iterations (PARSLR Full – blue squares). Methods (iii) and (iv) agree with the true value (True, black solid line).

6.9.1.2.2 Partial segmentation, hippocampus only

In the absence of a complete and accurate segmentation, as summarised in Table 6.2, the method of Rousset has been tested under two extreme conditions. When the remaining voxels were set as individual regions, the size of the $A$ matrix increased to the point that the least squared problem became ill-posed. The hippocampus was heavily underestimated with an average value of 3.84 for Rousset_vox (Figure 6.8, red circles) across the simulated noise levels. On the other hand, when all those voxels were set as a single region and assumed homogeneous, the result was still an underestimation that was very close to the uncorrected values. PARS LR (PARSLR_vox, blue squares – Figure 6.8) produced the best results with the hippocampus as the only segmentable region as the corrected values were the closest to the true value. It however seemed to go down in performance with increasing noise levels, but was still superior to all other methods being
assessed. This is further confirmed in the bar-chart in Figure 6.9 for MSE values where the MSE of PARSRL_vox under is better than LR and Rousset_back. In Figure 6.9, while the MSE values for PARSRL and Rousset under a complete and accurate segmentation is comparable (PARSLR_complete, Rousset_complete), Rousset produces substantially higher MSE values than PARSRL_vox for cases of Rousset where the remaining voxels are set as individual regions or a single background. The MSE for Rousset_vox is not shown in the figure as it is considerably higher (and thus out of scale on Figure 6.9) than the MSEs for the other methods with an average of 1.29 across the simulated noise levels.

**Figure 6.8:** Graph of mean intensity of Hippocampus versus different Poisson random noise levels ranging from High (COV = 144.2%) to Low (COV = 4.5%) under partial segmentation for (i) uncorrected hippocampus (green triangles), (ii) corrected with LR, 50 iterations (LR - black inverted triangles), (iii) corrected with Rousset’s method assuming remaining voxels as individual regions (Rousset_vox – red circles), (iv) corrected with Rousset’s method assuming remaining voxels set as a single background region (Rousset back – purple circles) and (v) corrected with PARSRL, 50 iterations, with remaining voxels set as individual regions (PARSLR_vox – blue squares).
Figure 6.9: Bar-chart for MSE of hippocampus across different levels of simulated Poisson noise from High (COV = 144.2%) to Low (COV = 4.5%) for (i) uncorrected data (green), (ii) Rousset under complete and accurate segmentation (Rousset_full - yellow), (iii) 50 iterations of PARSRL under complete and accurate segmentation (PARSLR_full - blue), (iv) 50 iterations of PARSRL under partial segmentation with only the hippocampus as a defined region and remaining voxels set as individual regions (PARSLR_vox, purple), (v) Rousset under partial segmentation with only the hippocampus as a defined region and remaining voxels set as background (Rousset_back – Red), (vi) 50 iterations of LR (grey). The MSE for Rousset_vox is not shown in the figure as it is considerably higher (and thus out of scale than the MSEs for the other methods) with an average of 1.29 across the simulated noise levels.

6.9.2 Measured data – 2D brain phantom

6.9.2.1 PVC Methods applied to Region1 using OP-OSEM data

When measured off the well-counter and corrected for decay and background counts, the mean activity concentration from Region1 was found to be 1710 kBq/mm². Figure 6.10 shows the calibrated results for the mean activity concentration in Region1, where the data reconstructed with OP-OSEM is compared with that from data reconstructed with OP-OSEM-PSF as well as those recovered from OP-OSEM with the methods of Rousset, PARSRL and LR under the philosophies described in Table 6.2.
Figure 6.10: Graph of mean activity concentration in region1 versus order of frames starting from Low stats (Average number of trues = 3.82x10^6) to High Stats (Number of trues = 2842.5x10^6) for (i) OP-OSEM (green triangles), (ii) OP-OSEM-PSF (cyan diamonds), (iii) LR (magenta crosses), (iv) Rousset_complete: Rousset under complete segmentation (red squares), (v) PARSLR_complete: PARSLR under complete segmentation (blue squares), (vi) Rousset_vox: Rousset with region1 under partial segmentation and the remaining pixels set as individual regions (purple left triangles), (vii) Rousset_back: Rousset with region1 under partial segmentation and the remaining pixels set as a homogeneous background (purple right triangles), (viii) PARSLR_vox: PARSLR with region1 under partial segmentation and the remaining pixels set as individual regions (black circles).

It is observed that under a complete segmentation, the methods of Rousset and PARSLR perform best. Most unusually, PARSLR was found to perform better than Rousset under a complete segmentation. This is because the scalp region also contained undefined pixels (i.e., pixels with no defined regions). These were kept as undefined within Rousset’s algorithm. However, in PARSLR any undefined pixel is assumed an individual region so that the results of PVC using PARSLR assume knowledge of these pixels as being defined regions. The amount of information that PARSLR interprets is more than what is / can be given to Rousset’s method. It is therefore potentially more accurate, even under complete segmentation for both methods.

LR and OP-OSEM-PSF have comparable values across the different levels of noise with a mean activity concentration of 1467±34 kBq/mm² for LR and 1467±32 kBq/mm² for OP-OSEM-PSF. PARSLR performs better than OP-OSEM-PSF across all noise levels.
When only Region1 is deemed segmentable and the remaining pixels set as individual regions, Rousset_vox (left-handed purple triangles) performed better than expected, delivering recovered mean intensities that were comparable to those recovered using PARSLR_vox (black circles). When all those remaining pixels were set as one homogeneous background, the recovered values (right-handed purple regions) were only slightly lower than those for PARSLR_vox.

### 6.9.2.2 PVC Methods applied to Region2 using OP-OSEM data

Region2 is effectively a CSF region where the activity concentration is zero.

![Graph of mean activity concentration in CSF region](image)

**Figure 6.11:** Graph of mean activity concentration in Region2 versus order of frames starting from Low stats (Average number of trues = 3.82x10^6) to High Stats (Number of trues = 2842.5x10^6) for (i) OP-OSEM (green triangles), (ii) OP-OSEM-PSF (cyan diamonds), (iii) LR (magenta crosses), (iv) Rousset_complete: Rousset under complete segmentation (red squares), (v) PARSLR_complete: PARSLR under complete segmentation (blue squares), (vi) Rousset_vox: Rousset with region1 under partial segmentation and the remaining pixels set as individual regions (purple left triangles), (vii) Rousset_back: Rousset with region1 under partial segmentation and the remaining pixels set as a homogeneous background (purple right triangles), (viii) PARSLR_vox: PARSLR with region1 under partial segmentation and the remaining pixels set as individual regions (black circles)

Figure 6.11 again illustrates the observation that PARSLR and Rousset’s method perform best under a complete and accurate segmentation. PARSLR_complete (blue squares) performs better than Rousset_complete (red squares) potentially due to the extra
amount of information that it interprets. When only Region2 is segmentable and the remaining pixels set as individual regions, PARSLR (PARSLR_vox – black circles) performs better than Rousset (Rousset_vox – left-handed purple triangles). In this case, when all those pixels are set as one homogeneous region, Rousset_back (right-handed purple triangles) compares to Rousset_vox.

6.10 Discussion

One of the observations from chapter 5 is that post-reconstruction methods for PVC like that of (Rousset et al. 1998) depend on the availability of a complete and accurate segmentation within at least a sub-volume of the image, if not of the whole image volume under investigation. With increasing heterogeneity, the number of regions within the image volume/sub-volume causes the A matrix of Rousset to grow larger and larger so that its inversion becomes an ill-posed problem. To understand why it turns into an ill-posed problem, consider two adjacent regions that are small compared to the PSF blurring. The blurred version of those regions will become quite similar such that the distinction of these 2 adjacent regions as column vectors in the A matrix is difficult. The matrix therefore becomes ill-conditioned. It would eventually create a solution dominated by noise that would spill into the ROI, introducing a bias within. In the limit that all voxels within the volume are defined as individual regions, the problem is a deconvolution that is solved using the LR iterative deconvolution under an EM framework. While it solves the same linear problem, its slow convergence and premature stopping makes it less prone to the ill-conditioning issue. It also gets rid of the requirement for accurately defining regions of interest, but is prone to Gibb’s artefacts at the borders of homogeneous regions within the image volume.

PARSLR was therefore developed to work as a hybrid between the region-based method of Rousset and the voxel-based method of LR: assumptions of homogeneity are allowed to be made where possible, with the algorithm running a voxel-based deconvolution where no regions can be accurately defined. The experiments described in section 6.8.2 under the philosophy summarised in Table 6.2 were carried out to assess the
performance of PARSLR with respect to the methods of Rousset and LR under conditions where these algorithms are known to break down.

6.10.1 Simulated hippocampus phantom

6.10.1.1 Complete and Accurate Segmentation

Under a complete and accurate segmentation, starting with noiseless data, both PARSLR and Rousset recovered the exact true value of the hippocampus showing concordance between methods. The presence of Gibb’s artefacts on the borders of homogeneous regions has also been confirmed from figure 6.6c.

With increasing levels of correlated noise, Rousset and PARSLR are comparable when a complete segmentation is known. The statistical significance of the better performance of Rousset_complete with respect to PARSLR_complete at the highest level of noise was tested using a paired-t-test for the difference between the 2 methods at 95% confidence interval and was found to be not significant (p=0.419 on two-tailed t-test) showing that the result has occurred more by chance due to the high level of noise in the data. The generally good performance of Rousset_complete under different levels of noise can be explained with reference to the works of (Aston et al. 2002) that points out that the noise model is likely to have an impact on the associated variance of the regions but not on the mean of the region which itself will remain unbiased.

It is also important to note that both LR and PARSLR have been presented under an EM framework which is known to be a slow converging algorithm. In our experiments, both LR and PARSLR have been stopped prematurely when the difference between successive updates of the mean intensity of the recovered region of interest were less than 1x10^-4. Since this occurred on average after 50 iterations, this was the number of fixed iterations used throughout the assessment. PARSLR and LR have both probably not converged completely in the presence of a higher level of noise. However, PARSLR was observed to converge faster than LR, when observed across the different noise levels as presented in figure 6.12 for the highest and lowest level of added correlated noise.
Figure 6.12: Rate of convergence of LR and PARS LR for low noise (blue) and high noise (red) with pure LR shown in dotted lines, PARS LR complete shown in full lines and PARS LR vox shown with dashed lines. Figure shows that PARS LR complete has the steepest curve, indicating a faster convergence to the true value which is equal to 5. PARS LR vox is then followed by LR.

6.10.1.2 Partial Segmentation (Hippocampus only)

In the event that only the hippocampus is accurately segmentable from this volume, 2 extremities of possibilities have been considered: either every remaining voxel is defined as an individual region or all the remaining voxels of the image volume/sub-volume is defined as one homogeneous background.

In the case of the former, as expected for Rousset’s method, the increase in the number of regions and therefore size of the $A$ matrix lead to its ill-posedness. More precisely, the inversion process resulted in the individual regions to become very noisy. These noisy voxels spilled back into the hippocampus causing the bias observed in its measurement (Rousset_vox; red circles in Figure 6.8)

In the latter case, Rousset_back eliminates the ill-posed condition of the inversion problem. However, with the neighbourhood of the hippocampus being effectively heterogeneous and the segmentation therefore inaccurate, the recovered mean intensity of the hippocampus is flawed.
PARSLR, on the other hand, is stable to leaving the unsegmented voxels as individual regions as it undergoes conventional LR iterative deconvolution over those voxels under an EM framework due to the nature of which allows it to lethargically progress to the optimum result. This explains why, in Figure 6.8, PARSLR (PARSLR_vox; blue squares) outputs the best recoveries for the mean intensity of the hippocampus over the different levels of noise simulated. The subsequent MSE scores of Figure 6.9 confirmed the agreement of PARSLR with Rousset under a complete and accurate segmentation as the bar charts for both Rousset_complete and PARSLR_complete lie very close to each other. Despite the MSE for LR being higher than those of Rousset_complete and PARSLR_complete for high count frames, the MSE is stable as the amount of noise increases. When only the hippocampus can be accurately segmented, PARSLR_vox (purple bar) delivers MSE values that are lower than Rousset_vox (results not shown because it is out-of-scale) and Rousset_back (red bars). MSE for PARSLR_vox was however observed to increase with increasing levels of noise. This increase can, in theory, be due to 2 reasons: either because the problem being solved is highly ill-conditioned such that it would be expected that the solution would contain noise propagated through the deconvolution process; or more simply because as the data gets noisier, it takes the algorithm more time to converge to the optimal solution and as the algorithm has been prematurely stopped at 50 iterations, this might be the reason for the increasing MSEs. Yet, PARSLR still manages to output the best corrected value for the mean intensity of the hippocampus under such conditions compared to the other methods under assessment.

6.10.2 Measured Data – 2D brain phantom

When the performance assessment was extended to measured data, the same trend observed with the hippocampus phantom was expected; that is, for Rousset_vox to fall apart and for PARSLR to sustain its stability when the remaining pixels were left as individual regions. For both Region1 and Region2, the performance of Rousset_vox was comparable to that of PARSLR_vox across all noise levels, suggesting that the $A$ matrix was invertible despite its size. There are 2 reasons why a square matrix cannot be inverted.
The first reason is obviously hardware where we are limited to the computational power available. For example, if a computer has 1GB of RAM dedicated for the computation of the inverse of the $A$ matrix, the maximum size of $A$ that could be dealt with is $\frac{1024^3}{8} \approx 11000^2$ meaning that only a maximum of 11,000 individual regions can be defined (assuming computation in Matlab with each number being an 8-byte double floating point).

More importantly however, whether the square matrix is invertible or not depends on how well-conditioned the matrix is. The condition number of a matrix, under the Euclidean norm and associated matrix norm, is the ratio of the largest singular value to the smallest. A condition number close to 1 would imply a well-conditioned matrix where the inverse can be obtained with good accuracy. The larger the condition number, the more ill-conditioned the matrix. If a plot of singular values is made against the size of a matrix, the faster the singular values fall to zero, the larger will be the condition number, and hence, the worse the accuracy.

The difference between the $A$ matrix used in the hippocampus phantom and that used in the measured data is that it was obtained with 2 different blurring kernels. For the hippocampus phantom, a Gaussian kernel of 2.5mm FWHM was used. For the measured data, the kernel recommended for the HRRT (Comtat et al. 2008) was used. We explored the implication of using different kernels for calculation of the $A$ matrix for the measured data. For the same matrix size, we generated the $A$ matrix by blurring with different kernel sizes and used the resulting $A$ matrix to calculate the corrected value for Region1.

We plotted the singular values for each generated A-matrix and compared the PVC results obtained between the methods of Rousset_vox and PARSLR_vox when using those selected blurring kernels.

From figure 6.13, which is a plot of the singular values within the $A$ matrix versus the order of the number of regions (from highest to lowest) within the $A$ matrix for blurring kernels of different sizes, it is clear that the PSF model, as recommended by (Comtat et al. 2008), is the best conditioned among the blurring kernels explored. The singular values
(solid black line) have not fallen off as fast as the other kernels’ have. The Gaussian kernels were less conditioned as compared to the one by Comtat et al.. As the size of the kernels increased, the singular values fell off even faster.

The observation of the rate of fall of the singular values corroborates with the partial volume corrected values obtained from the $A$ matrix generated from these kernels for Rousset_vox. Figure 6.14 represents the results of applying Rousset_vox and PARSLR_vox assuming those blurring kernels. It can be observed that Rousset_vox produces a corrected image where the remaining pixels are seemingly not hampered by noise (invisible to the naked eye, but there are some pixels that have fallen apart). As the kernel is moved from that recommended by (Comtat et al. 2008) to Gaussian kernels, the remaining pixels get noisier confirming that the singular values are falling at a steeper rate thus undermining the accuracy of the inversion operation. In order to understand the reason why the kernel recommended by Comtat was performing better, the latter was modified to model a steeper rate of fall of the singular values. This was achieved by using a wider head and narrower tail in the 2 Gaussians (input details). It can be observed that the singular values for the modified Comtat kernel followed that of a 2.5mm Gaussian kernel. The result of using this modified kernel (1585 kBq/mm$^2$) was comparable to those obtained with the 2.5mm Gaussian kernel (1577 kBq/mm$^2$).

An interesting observation from Figure 6.14 is that as we head further and further away from a good estimation of the blurring kernel, the method of Rousset tends to fall apart quicker and in a more drastic fashion than PARSLR. In the latter, as the blurring kernel is overestimated, the algorithm will try to recover a lower number of frequencies than with the correct blurring kernel hence introducing the bias in the measurement. In Rousset’s method, the blurring kernel is used to calculate the spread of one region into its neighbours. With a larger kernel, the fall in amplitude of the central pixel gets more drastic and thus has an impact on the singular values for the $A$ matrix. The latter becomes quickly ill-conditioned such that the inversion is no more accurate. This finding finds its importance when estimating the blurring kernel to be used on imaging data that has been further processed post-reconstruction. For example, if image-based motion correction is applied to the reconstructed data, the final step of the motion correction process, as seen
in Chapter 4, involves back-projecting all the frames into a single reference position. Moving and saving the image into this reference position requires interpolation of the data points, which is known to cause a general smoothing of the image. Estimating the new blurring kernel is important for accurate recovery values. However, the blurring kernel seems to be more of an issue with Rousset’s method than with PARSLR.

Figure 6.13: Graph of singular values for blurring kernels of different sizes presented on log scales and shows that the smallest singular value for Comtat’s kernel (black plot) is bigger than those of the Gaussian kernels.

![Graph of singular values for Gaussian kernels of varying widths](image)

Figure 6.14: Recovered values for Region1 with PARSLR and Rousset using different blurring kernels.
When the remaining pixels were coalesced into a single homogeneous background, the performance of Rousset_back for Region1 was comparable to Rousset_complete for Region1 (Figure 6.10) but only comparable to Rousset_vox for Region2 (figure 6.11). The clear under-performance of Rousset_back as observed with the hippocampus phantom is not observed here. The reason for this is that despite extending the background region in an attempt to make it more heterogeneous, the average value of the latter was still close enough to the value of the surroundings that contained significant PSF values; that is, Region1 had a background whose average was comparable to the average value of WM around it while Region2 had a background with an average value comparable to GM. Rousset_back therefore produced corrected values with only slight discrepancies.

6.10.2.1 Comparison with OSEM-PSF data

There has been few works presented in literature regarding the resolution properties of the EM algorithm that includes PSF modelling (Thielemans et al. 2010), as opposed to its widely documented noise properties (Barrett et al. 1994). Image reconstruction is an ill-conditioned problem and typically, EM algorithms, such as OSEM, are stopped prematurely before they reach the maximum-likelihood solution, where the development of higher frequencies lead to noisy estimates (Angelis et al. 2011). The EM algorithm has 2 stages: (i) the early stage where the updates are large and fast as the lower frequency components are being recovered; (ii) the later stage where the updates are smaller and slower as the higher frequency components are recovered. If the PSF model is included in the reconstruction, the updates are smoother and take a longer time to converge as compared to OP-OSEM without PSF. For this reason, more tomographic iterations are performed on OP-OSEM-PSF (12 iterations) than OP-OSEM (10 iterations). In total, it can be said that OP-OSEM-PSF has undergone 12x16 = 192 tomographic iterations (because of 16 subsets) in addition to 192 LR iterations for the PSF modelling and explains why the reconstruction that includes the resolution model is better than that without.

However, when 50 iterations of LR are applied to OP-OSEM data, the result becomes akin to that of OP-OSEM-PSF. Despite a lower number of iterations (50 compared to 192), the results between LR and OSEM-PSF are very close as the LR algorithm under an EM framework progresses very slowly towards the maximum-likelihood solution.
PARSLR offers an improved quantified value as compared to OP-OSEM-PSF and LR as it contains \textit{a priori} information on homogeneous regions that exist within the volume. As seen in figure 6.12, this tends to accelerate the converge of the algorithm towards the maximum-likelihood solution.

\section*{6.11 Conclusions}

In this chapter, we have designed and implemented a novel post-reconstruction partial volume correction algorithm that addresses the weaknesses of currently used algorithms in clinical research. Unlike Rousset’s method, PARSLR does not require a complete and accurate segmentation of the image volume or sub-volume. It only requires an accurate segmentation of the volume of interest and performs a region-based PVC there. Outside the identified region, it performs a voxel-based deconvolution which gets rid of the ill-posed problem faced by Rousset’s method when faced with a matrix too large to invert. In PARSLR, the Gibb’s artefacts on the borders of homogeneous regions are suppressed. They cannot exist due to the homogeneity constraint.

We have successfully demonstrated, via the use of a simulated hippocampus phantom; and measured data, via a physical brain phantom, the conditions under which Rousset’s method falls apart and the ability of PARSLR to better recover the region of interest. We have also been able to demonstrate that the estimation of the correct blurring kernel is very important for post-reconstruction PVC as it seems to dictate how well-conditioned the problem is.. While the limit of the blurring kernel recommended by (Comtat et al. 2008) could not be explored due to computational limitations, we intuitively know that it ultimately falls apart well under the size of the imaging volume itself. PARSLR always handles the whole image volume for the same computational power.

The noise model applied to the simulations was a purely correlated noise model. (Aston et al. 2002) have widely discussed the difficulty in producing a correct noise model. Doing so was beyond the scope of this work which was only to look into what happens to region-based methods when the important homogeneity assumption breaks down.
It is believed that method finds its importance in clinical research in many cases. For example, where it is required to measure a region that is in proximity of a lesion or structure that prevents a complete and accurate segmentation or for in an epilepsy study using $[^{11}\text{C}]-\text{Verapamil}$ where the hot choroid plexus spills into the hippocampus. The accuracy of PVC results using Rousset’s method is difficult to gauge if the heterogeneity of the choroid plexus is considered.

Finally, in a dementia context, with brain images suffering from enlarged ventricles and therefore CSF spaces, and automatic segmentation of the Medial Temporal Lobe being inherently difficult, being able to segment out (manually) the hippocampus and leave the remaining voxels as individual regions for PVC is a robust alternative to estimate more accurately the activity concentration within this region of interest.
Chapter 7  
Summary of outcomes and future prospects

7.1 Summary of thesis

In this thesis, we have underlined the importance of early and accurate diagnosis of Alzheimer’s Disease and the useful role of PET as a medical imaging technique that allows investigation of AD’s pathology in vivo (Chapter 1). The high resolution of PET was explained (Chapter 2) and has the potential to address unanswered clinical questions like the separation of atrophy and glucose metabolism in brain tissues in AD, which involves the accurate quantification of atrophied small structures.

While the high resolution problem in brain PET has been addressed with the development of a dedicated high resolution brain PET scanner, the HRRT, access to this high resolution is limited by (i) movements and (ii) partial volume effects which, though lesser than as compared to in other clinical scanners, is still present.

The objectives of this work were therefore to develop methods for motion correction and Partial Volume Correction to allow access to the benefits of PET imaging at high resolution.

With a high resolution scanner also comes the need for motion correction, a review of which is given in chapter 3. The review focused on brain imaging only and highlighted the general agreement that event-based motion correction is better than image-based motion correction, but requires not only external detection of movements which is often undesired in a clinical setting (like dementia) but also accurate quantification which has been questioned for existing systems. The practicality of acquiring and storing data in listmode format was also discussed, at the end of which 2 hypotheses were formulated:

(i) Whether raw listmode data could be used to detect when movements occur.
(ii) Whether, in practice, subsequent reframing of the listmode data based on those detected movements reduces degradation of the data.
Chapter 4 then details the development of an improved frame-by-frame realignment algorithm (RAS_FBF). The centroid of activity during the scan is calculated from the raw listmode data and based on centroid analysis, the listmode data is partitioned into shorter frames that have minimal motion within. The ability of the centroid analysis to detect when movements occur versus its ability to accurately quantify those movements is also looked into via moving a cylindrical phantom within the scanner’s FOV using the movable bed and comparing the changes in centroid coordinates with the truth obtained from the bed position as well as the measurements obtained from co-registration. The observation was then extended to clinical data, comparing coordinates of the centroid obtained via calculation from listmode versus those obtained via co-registration to the reference frame. Finally, the added benefits of reframing data according to where movements were detected were assessed by comparing RAS_FBF with conventional frame-by-frame realignment (FBF) post-reconstruction and Attenuation and Scatter corrected frame-by-frame realignment (AS_FBF) for 6 patients that formed part of an AD study using FDG-PET.

The partial volume problem that is inherently due to the lower intrinsic resolution of PET scanners is reviewed in chapter 5. The object-image relationship is derived and a brief overview of partial volume correction strategies in literature is given. 2 groups of methods are explored in further details. The first division is the separation of methods that use image-based segmentation and the other that performs image-based deconvolution to recover resolution. We establish how, in either case, the problem is effectively a linear one, where for the former case, can be solved using different matrix inversions and linear least squared methods, depending on its formulation that is, in turn, based on the choice of partitions in both the true and observed image. The weakness of such region-based methods is highlighted as they require a complete and accurate segmentation at least within a sub-volume of the image, if not the whole image volume. As the level of heterogeneity increases, the matrix to invert becomes larger and its inversion becomes an ill-posed problem. In the limit that every voxel is an individual region, the problem becomes a deconvolution one that is observed to be best solved using the Lucy-Richardson iterative deconvolution algorithm. However, the latter is prone to Gibb’s artefacts on the borders of homogeneous regions that is a consequence of the
deconvolution problem where lost high frequency terms cannot be recovered. At the end of this review chapter, the hypothesis formulated is whether it is possible to modify the LR algorithm presented under an EM framework to include assumptions regarding homogeneity of one or more regions within an image volume and leave the remaining unsegmentable regions as voxels to undergo a conventional LR iterative deconvolution.

In chapter 6, the novel method (PARSLR) that works as a hybrid between region-based and voxel-based method for partial volume correction is developed. The problem is formulated using LR iterative deconvolution under an EM framework and the algorithm modified to include information about homogeneity within regions. The algorithm is presented under an efficient implementation that allows it to be applied over the whole image volume. The ability of PARSLR to perform better than the popular region-based method of Rousset (Rousset et al. 1998) in an absence of complete and accurate segmentation, while suppressing the Gibb’s artefacts observed on the voxel-based LR iterative deconvolution was assessed.

A 3D hippocampus phantom was created and tested for accuracy of recovered values under a complete segmentation, and under partial segmentation considering the 2 extreme cases: (i) when the remaining voxels are defined as individual regions (PARSLR_vox versus Rousset_vox) and (ii) when the region of significant PSF values outside the hippocampus is considered a single homogeneous background (PARSLR_vox versus Rousset_back). The experiment was extended to measured data in a 2D brain slice phantom printed with radioactive fluorine ink and 2 selected regions corrected for partial volume under the same philosophy.

**7.2 Research findings and extrapolation to the medical context**

Two methods of data corrections have been successfully developed in a bid to improve the resolution of the PET data and accuracy of quantitative measurements in PET.

In chapter 4, the performance of RAS_FBF was compared to FBF and AS_FBF on 6 FDG-PET scans. In a visual assessment by 4 independent observers, RAS_FBF was found to offer images with highest quality. In quantitative evaluation via Time-Activity Curves,
RAS_FBF was found to deliver the biologically most plausible activity concentrations for selected regions as compared to FBF and AS_FBF.

Unsurprisingly, the benefits of RAS_FBF over AS_FBF were found to be dependent on when motion occurred and the type of motion that occurred. RAS_FBF is particularly effective in addressing movements detected within-frames. Where these movements were of no big consequence, qualitative and quantitative differences between RAS_FBF and AS_FBF were still observed suggesting benefits of reframing the data during motion correction. In fact, the performance of AS_FBF exceeded expectations probably because of the good \textit{a priori} frame definitions for dynamic data acquisition in a clinical setting. Importantly, RAS_FBF was never worse than AS_FBF. The centroid analysis technique is effective at detecting discrete movements and slow drifts but cannot separate rapid intra-frame movements like tremors or motion due to respiratory and cardiac cycles.

The RAS_FBF method will enable PET measurements with improved decoupling from motion (it will not completely remove motion as some types of motion are not addressed). The effect of motion has two potential degrading effects for clinical studies: 1) it introduces variability into the measurements meaning that more subjects need to be studied; 2) it could result in false positive if there is a correlation between motion and condition being studied. Taking the case of the hippocampus and putting it into context, the region of interest is surrounded by enlarged CSF spaces, as well as white matter tracts and other regions of the medial temporal lobe. It is observed that with RAS_FBF, the activity is higher than AS_FBF, meaning that in a group study, the metabolism for hippocampus in AD can be substantially higher. The benefit of RAS_FBF is that it will now allow clinicians to evaluate whether the disease’s progression from normal or MCI to AD is effectively accompanied by a decrease in metabolic activity within the hippocampus or whether the latter is preserved, the condition of changes due to movement having been minimised to the best of abilities.

The development of PARSLR (chapter 6) has the same clinical objective as the development of RAS_FBF, even if the inherent reasons as to why motion correction and partial volume correction are applied are different. Motion is a subject-related problem that degrades resolution as well as reduces the accuracy of quantitative measurements. It
is variable depending on the clinical conditions of the patients. Partial volume problems on the other hand, are due to the scanner’s limited resolution and therefore a physical problem. However, it is desired to know whether an atrophied region is effectively experiencing decreased metabolism or whether the smaller sampled region is experiencing partial volume effects that have not been properly recovered. Again the hippocampus is considered, which is located in an area of the brain where automatic segmentation is inherently difficult, besides the problem of the presence of enlarged CSF spaces and activity from the choroid plexus which is known to be a very vascular region. Finding an accurate measurement of the activity concentration within the hippocampus is difficult as assuming a complete and accurate segmentation is as difficult, if not impossible. PARSLR allows one to accurately delineate only the hippocampus and leave the remaining regions as voxels for PVC.

With the simulated 3D hippocampus phantom, when only a partial segmentation is available, application of Rousset’s method experienced an ill-posed condition as the matrix for inversion became too large such that it introduced an error in the measurement of the hippocampal intensity. PARSLR performed well irrespective of whether a complete and accurate, or partial segmentation was available. It was also shown that one solution to the problem with Rousset and too many regions is to approximate a heterogeneous region for which accurate segmentation is problematic with a homogeneous region. As was demonstrated such an approach can lead to significant errors in the region of interest. Under a complete and accurate segmentation, all methods were in agreement.

In extending the experiments to measured data, same results as with the 3D hippocampus phantom were not observed. One possible reason for this is that the Comtat kernel offered better conditioning of the $A$ matrix. One possible conclusion from this is the potential importance of making sure that the correct kernel is used for recovering the true value as a slight change in the kernel can cause a marked discrepancy in the corrected value. More work has to be done on this front to confirm such a hypothesis though.
Image reconstruction with PSF modelling produced similar results as compared to application of LR deconvolution post-reconstruction on data reconstructed without PSF modelling. This shows that performing the deconvolution step during or after the reconstruction is likely to yield the same results. PARSLR performed better than both LR and OSEM with resolution modelling due to the presence of additional information that has potentially accelerated the convergence of the EM algorithm.

7.3 Future work

7.3.1 Clinical application for RAS_FBF

So far, the performance of the proposed RAS_FBF method has demonstrated qualitative and quantitative improvements on patient data but has not been used to look into the outcome of a clinical investigation.

The clinical importance of RAS_FBF can be further shown by comparing uncorrected data with motion corrected data between a group of patients that have a baseline scan and a follow-up scan with noticeable cognitive decline. 3 regions are selected, one which is not affected by the disease’s pathology (example pre-central gyrus) and one where we expect to see a marked change (example angular gyrus) and one where the changes in metabolism is unknown (hippocampus). It would then be expected that the regional concentration in activity between baseline and follow-up scans for the pre-central gyrus would remain the same while there will be noticeable change in a region like the angular gyrus. This would then increase the confidence in findings in changes for the activity concentration of the hippocampus across the group.

7.3.2 Further investigations for PARSLR

A region like the hippocampus can be manually segmented and PARSLR applied for PVC without requirement of structural information for the surroundings. Potential further investigations can then be lead on 2 fronts:
(i) With an accurately segmented hippocampus (manual), application of PARSLR to clinical data and comparison of performance of PARSLR against Rousset and LR with the observations from chapter 6. Are the same trends observed?

(ii) Use of PARSLR and manually segmented region as a gold standard and observe the results with existing algorithms assuming a finer segmentation that is within the limits of a conditioned problem (example, assuming choroid plexus as a homogeneous region, lateral ventricles as another, CSF spaces as another, assuming heterogeneous white matter within the medial temporal lobe etc)

7.3.3 Combining both methods for group studies in clinical research

Chapter 1 has exposed the difficulties in being able to apply robust data correction techniques to clinical PET data, which is why state of the art research facilities have reported different findings in the relationship between atrophy and metabolism in a key region like the hippocampus (Herholz et al. 2006; Samuraki et al. 2007; Chetelat et al. 2008; Mistur et al. 2009; Mosconi et al. 2009). One of the observations was that data correction techniques like motion and partial volume correction were either not applied or applied under wrong assumptions (like complete segmentation for example).

Now however, methods for both motion correction and partial volume correction have been established and have been shown to be robust and reliable in a clinical setting. The main advantage of these methods is that they can be retrospectively applied to clinical data as the RAS_FBF does not require any external equipment for motion detection and relies only on listmode data acquisition and PARSLR works on post-reconstructed data. Subsequent application of both data correction methods will produce data which is clinically more accurate and help in a more precise understanding of the disease’s pathology.


Jin, X., Sandiego, C. M., et al. (2010). "Multiple Acquisition Frame-Based Motion Correction for Awake Monkey PET Imaging.", Medical Imaging Conference, Conference Record, 2010 IEEE


