Post-processing Analysis for Magnetic Resonance Renography

D. Rodriguez Gutierrez

Submitted for the Degree of Doctor of Philosophy from the University of Surrey

Centre for Vision, Speech and Signal Processing
School of Electronics and Physical Sciences
University of Surrey
Guildford, Surrey GU2 7XH, U.K.

March 2008

© D. Rodriguez Gutierrez 2008
Summary

Functional studies are routinely performed to provide quantitative evaluations of renal function. Of these, Nuclear Radioisotope Renography is the current technology of choice in clinical studies. However, there has been considerable interest in using Magnetic Resonance (MR) in these studies, motivated by avoidance of the use of ionising radiation, which is particularly attractive for paediatric subjects and by the availability of high quality anatomical and functional information within a single investigation, thus allowing simultaneous observation of pathology and renal function. The overarching factor preventing widespread use of MR Renography is the lack of image processing techniques capable of robust extraction of quantitative information from the available data. The work presented in this thesis is aimed at addressing this issue.

A methodology for estimating renal function is proposed. Movement correction of Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) data was performed using a 3D movement correction method based on the phase difference of time-adjacent data volumes. This was followed by a novel approach demonstrating the first Partial Volume (PV) Effect correction on MR Renography data. For a typical renal Region of Interest, the observed intensity for each voxel was de-composed into its constituent parts, corresponding to the contributions from each tissue, using knowledge of the Point Spread Function (PSF) and high-resolution registered templates for each anatomical tissue type. Thus, non-renal contributions from liver, spleen and other surrounding tissues could be eliminated from time-intensity curves. This produced a change the renal curve that resulted in enhanced Glomerular Filtration Rate (GFR) estimates, as per a Rutland-Patlak analysis of the time intensity data: a cohort of 10 healthy volunteers produced a mean enhancement of 36% in relative GFR with a mean improvement of 5% in $R^2$ fitting the Rutland-Patlak model when PV correction was applied compared to no PV correction.

PV correction in small structures (with respect to the PSF) corresponding to arterial vasculature were also investigated using a statistical Partial Volume classifier and a novel mixing prior that models more closely asymmetries in intensity distributions related to small object size than others proposed in the literature. Classifier performance was found to be approximately inversely proportional to Contrast to Noise. An intensity recovery step was also proposed for Arterial Input Functions derived from such small objects. The resolution of the PSF of the DCE-MRI sequence was found to similar to typical aortic diameters and therefore would present intensity reductions due to the PV Effect of $\approx 60\%$ that need to be recovered.

This work has shown that in addition to movement correction, contrast agent quantification and accurate modelling, absolute GFR measurement via DCE-MRI is significantly affected by the PV Effect, and as such must be robustly accounted for.

Key words: Magnetic Resonance Renography, Partial Volumes, Renal Function.
Acknowledgements

First of all, I would like to thank my supervisor, Kevin Wells, for his guidance and help. I particularly appreciate his patience, encouragement, and all those corrections ... despite the copious amounts of red ink!

Secondly, I would like to thank the Centre for Vision, Speech and Signal Processing for both funding this research and providing an enjoyable and challenging environment to work. I would also like to thank the University of Surrey’s Faculty of Engineering and Physical Sciences, the UK Institute of Physics, and the IEEE for helping me financially to attend the MIC and ESMRMB conferences.

I would also like to thank Iosif Mendichovszky and Isky Gordon from the Institute of Child Health. They provided the data and clinical knowledge behind this work. It was a privilege of working with them. My thanks also go to everyone involved in the kidmri.org group from which I learn so much about our little kidneys. If MR Renography is ever used in routine clinical practice it will be greatly due to the work and knowledge contributed by the people in this group. There have been many others with whom I had the pleasure to collaborate during this time: John Chiverton, Beier Jia, Mike Partidge, Oliver Diaz Montesdeoca, Acerina Moran Santana, Haval Kadhem and Mary Yip. Thanks for all your help and I hope that we can carry on working together in the future.

I’ve spent quite a few years in Guildford, and it was not all work in the Medical Imaging office from where I have fond memories of many friends. Their help was as important as anything else in getting me through. In this time, many other friends have come and gone through CVSSP and the University. I thank you all and miss many of you that are no longer around.

Liliana, you know that the work presented here is very much thanks to you. It’s pretty hard to do a PhD, but coping with two at the same time really takes some extra patience. Thanks for being there, sometimes simply to listen, to give me that last push, or to organise my otherwise random thesis writing approach.

Finally, I would also like to thank my family. This work is the result of their unconditional love and support. The same that I had for Avelino, Isabel, and Juan, who were with us at the start of this PhD but are no longer. However, in their place, Isabella has arrived to rejoice all our hearts. Now that the PhD is finished I look forward to being with you more often and enjoying our time together ... even if it always rains in Asturias!
Abbreviations

ADH  Anti-Diuretic Hormone
AIF  Arterial Input Function
AUC  Area Under the Curve
BOLD Blood Oxygenation Level Dependent
CDF  Cumulative Density Function
CNR  Contrast to Noise Ratio
CT   Computed Tomography
DCE-MRI Dynamic Contrast Enhanced Magnetic Resonance Imaging
DRF  Differential Renal Function
DTPA Diethylene-Triamine-Penta-Acetic (molecule)
EBCT Electron Beam Computed Tomography
ERPF Effective Renal Plasma Flow
ESF  Edge Spread Function
FA   Factor Analysis
FISP Fast Imaging with Steady state Precession (sequence)
FLASH Fast Low Angle Shot (sequence)
FOV  Field of View
FSRG Feedback Shift Register Generator
GFSRG Generalised Feedback Shift Register Generator
FT   Fourier Transform
GE   Gradient Echo (sequence)
GFR  Glomerular Filtration Rate
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GT</td>
<td>Ground Truth</td>
</tr>
<tr>
<td>GTM</td>
<td>Geometric Transfer Matrix</td>
</tr>
<tr>
<td>ICG</td>
<td>Inverse Cumulative Gaussian</td>
</tr>
<tr>
<td>IFT</td>
<td>Inverse Fourier Transform</td>
</tr>
<tr>
<td>IG</td>
<td>Inverse Gaussian</td>
</tr>
<tr>
<td>IRSE</td>
<td>Inversion Recovery Spin Echo (sequence)</td>
</tr>
<tr>
<td>IVU</td>
<td>Intra-Venous Urogram</td>
</tr>
<tr>
<td>KUB</td>
<td>Kidney, Ureter and Bladder (X-ray)</td>
</tr>
<tr>
<td>LCG</td>
<td>Linear Congruent Generator</td>
</tr>
<tr>
<td>LFG</td>
<td>Lagged Fibonacci Generator</td>
</tr>
<tr>
<td>LSF</td>
<td>Line Spread Function</td>
</tr>
<tr>
<td>LTI</td>
<td>Linear-Time-Invariant (system)</td>
</tr>
<tr>
<td>MDRD</td>
<td>Modification of Diet in Renal Disease (formula)</td>
</tr>
<tr>
<td>MI</td>
<td>Mutual Information</td>
</tr>
<tr>
<td>MR</td>
<td>Magnetic Resonance</td>
</tr>
<tr>
<td>MTT</td>
<td>Mean Transit Time</td>
</tr>
<tr>
<td>NI</td>
<td>Nuclear Imaging</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>PAH</td>
<td>Para-AminoHippurate (acid)</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PET</td>
<td>Positron Emission Tomography</td>
</tr>
<tr>
<td>PRNG</td>
<td>Pseudo Random Number Generator</td>
</tr>
<tr>
<td>PSF</td>
<td>Point Spread Function</td>
</tr>
<tr>
<td>PV</td>
<td>Partial Volume (effect)</td>
</tr>
<tr>
<td>RBF</td>
<td>Renal Blood Flow</td>
</tr>
<tr>
<td>RC</td>
<td>Recovery Coefficient</td>
</tr>
<tr>
<td>PD</td>
<td>Proton Density</td>
</tr>
<tr>
<td>RF</td>
<td>Radio Frequency</td>
</tr>
<tr>
<td>ROI</td>
<td>Region Of Interest</td>
</tr>
</tbody>
</table>
Abbreviations

RPF  Renal Plasma Flow
RR   Radioisotope Renography
SE   Spin Echo (sequence)
SNR  Signal to Noise Ratio
SPECT Single Photon Emission Computed Tomography
SPGR Spoiled Gradient Recalled (sequence)
SPIO Super-Paramagnetic Iron Oxide (particles)
SVD  Singular Value Decomposition
TA   Time-activity (curve)
TE   Echo Time
TR   Repetition Time
VIBE Volumetric Interpolated Breath hold Examination (sequence)
Contents

Abbreviations vii
Contents xi
List of Figures xvii

1 Introduction 1
  1.1 The Kidneys ................................................................. 2
    1.1.1 Macroscopic Kidney Anatomy ................................. 2
  1.2 Renal Function ............................................................ 4
    1.2.1 Filtration ............................................................... 4
  1.3 Non Image-based Renal Function Assessment .................. 7
    1.3.1 Measuring Renal Plasma Flow and Renal Blood Flow ..... 8
    1.3.2 Measurement of Glomerular Filtration Rate .............. 9
  1.4 Biomedical Imaging and Magnetic Resonance Renography ... 11
    1.4.1 Renal Imaging ....................................................... 12
    1.4.2 Image Processing for MR Renography ..................... 14
  1.5 Aim and Objectives ...................................................... 16
  1.6 Contribution ............................................................... 16
  1.7 Thesis Overview .......................................................... 17

2 Image-based Renal Function Estimation 19
  2.1 Radionuclide Renography ............................................. 20
    2.1.1 Standard Renography ............................................ 21
  2.2 MR Renography .......................................................... 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.2.1</td>
<td>Avoiding the $T_2^*$-shortening Effect in the Kidney</td>
<td>28</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Flow Effect</td>
<td>29</td>
</tr>
<tr>
<td>2.3</td>
<td>Movement Correction</td>
<td>30</td>
</tr>
<tr>
<td>2.3.1</td>
<td>Manual Correction</td>
<td>32</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Hough Transform</td>
<td>32</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Distance Minimisation</td>
<td>33</td>
</tr>
<tr>
<td>2.3.4</td>
<td>Phase Difference</td>
<td>35</td>
</tr>
<tr>
<td>2.3.5</td>
<td>Assessing Accuracy</td>
<td>37</td>
</tr>
<tr>
<td>2.4</td>
<td>Kidney Segmentation</td>
<td>38</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Whole-Kidney/Parenchyma Segmentation of Functional Data</td>
<td>40</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Intra-kidney Segmentation of Functional Data</td>
<td>40</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Non-rigid Kidneys</td>
<td>42</td>
</tr>
<tr>
<td>2.5</td>
<td>Partial Volume Effect</td>
<td>43</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Template-based PV Classification</td>
<td>44</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Probabilistic PV Classification</td>
<td>46</td>
</tr>
<tr>
<td>2.6</td>
<td>Modelling Kidney Function</td>
<td>50</td>
</tr>
<tr>
<td>2.6.1</td>
<td>Multi-compartment Renal Models</td>
<td>51</td>
</tr>
<tr>
<td>2.6.2</td>
<td>Localised Tracer Kinetics in the Kidney</td>
<td>52</td>
</tr>
<tr>
<td>2.7</td>
<td>Conclusions</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>MR Data Acquisition and Preprocessing</td>
<td>57</td>
</tr>
<tr>
<td>3.1</td>
<td>Data Acquisition</td>
<td>57</td>
</tr>
<tr>
<td>3.2</td>
<td>Linearity of Signal Intensity vs. Gd Concentration</td>
<td>59</td>
</tr>
<tr>
<td>3.3</td>
<td>Determining the PSF of the Dynamic Data Sequence</td>
<td>59</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Point Spread Function</td>
<td>60</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Line Spread Function</td>
<td>62</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Line Spread Function Estimation</td>
<td>62</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Empirical LSF Model for the DCE-MRI sequence</td>
<td>67</td>
</tr>
<tr>
<td>3.4</td>
<td>Movement Correction</td>
<td>73</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Method</td>
<td>73</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Results and Discussion</td>
<td>76</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusions</td>
<td>77</td>
</tr>
</tbody>
</table>
# 4 Synthetic Data Simulator

4.1 Simulated Data in Medical Imaging ................................................. 80

4.1.1 Generation of Synthetic MR Data ........................................... 81

4.2 A Synthetic Data Generator ............................................................ 82

4.3 High Resolution Templates ................................................................. 83

4.3.1 3D Geometric Shapes .................................................................................. 84

4.3.2 Anatomical Data Segmentation .................................................... 86

4.4 Movement Model .................................................................................. 87

4.5 Tissue and Tracer Models ................................................................. 87

4.5.1 Tissue and Tracer Intensity Models ........................................... 88

4.5.2 Noise Models ......................................................................................... 88

4.5.3 Generation of Synthetic Gaussian Noise ........................................ 89

4.6 Downsampling ....................................................................................... 90

4.6.1 Low Pass Filtering ................................................................................ 91

4.6.2 Effect of Filtering on Variance ............................................................ 91

4.6.3 Decimating ............................................................................................. 93

4.7 Generated Synthetic Data ..................................................................... 93

4.7.1 Synthetic 4D DCE-MRI Kidney Data ........................................... 93

4.7.2 Synthetic 3D PET Phantom Data .................................................... 96

4.8 Conclusions ......................................................................................... 97

# 5 Partial Volume Correction of DCE-MRI Renal Data

5.1 Theory ..................................................................................................... 100

5.1.1 Construction of a PV Mixing Map ............................................... 100

5.2 Methodology ........................................................................................ 102

5.2.1 Estimation of the Mixing Maps ................................................... 103

5.2.2 Movement Correction and Template Registration ..................... 105

5.2.3 Extraction of Tissue Component Curves ....................................... 106

5.2.4 Extraction of Arterial Input Functions ........................................... 106

5.3 Results .................................................................................................... 107

5.3.1 Effect of Non-renal Contributions to Cortical ROIs .................... 107

5.3.2 Component curves ............................................................................ 109

5.3.3 Patlak Analysis .................................................................................... 111

5.4 Discussion ............................................................................................... 114

5.5 Conclusions ............................................................................................ 117
## Contents

D Generation of Synthetic Noise 157

D.1 Uniform Random Numbers .................................................. 158

D.2 Generating Other Probability Distributions ......................... 160

D.2.1 Direct Sampling ............................................................... 160

D.2.2 Rejection Method ............................................................ 162

D.2.3 Mixed Method ................................................................. 162

D.2.4 The Box-Muller Transform .............................................. 163

Bibliography 165
List of Figures

1.1 Cross-section Diagram of the Kidney ................................................................. 3
1.2 Cortical Tissue and the Glomerulus .................................................................... 5
1.3 Autoregulation of GFR ....................................................................................... 7

2.1 Example of a Typical Radionuclide Renography Image .................................... 21
2.2 Example of Typical Radionuclide Renography Renograms ............................. 22
2.3 Rutland-Patlak Model ....................................................................................... 26
2.4 Cortical and Medullary MR Renograms ........................................................... 27
2.5 Effects of Breathing Movement on DCE-MRI data ........................................ 31
2.6 The Partial Volume Effect ................................................................................. 44
2.7 PV Likelihood Models I .................................................................................... 47
2.8 PV Likelihood Models II ................................................................................... 48
2.9 Mixing Prior Functions ...................................................................................... 49
2.10 Cortical Compartment Model .......................................................................... 52
2.11 Gamma Variate Function ................................................................................. 53
2.12 Three-component Cortical Uptake Model ....................................................... 54

3.1 MR Renography Data Examples ...................................................................... 60
3.2 Relationship Between SI and GD Concentration ............................................. 61
3.3 MR Phantom ...................................................................................................... 63
3.4 Oversampling an Edge ...................................................................................... 64
3.5 Registered LSFs ................................................................................................. 65
3.6 Effects of Poor Sampling .................................................................................. 66
3.7 Effects of the Magnitude Operator .................................................................... 67
3.8 LSF Model Fitting .............................................................................................. 69
List of Figures

3.9 LSF Model Fitting (II) ............................................... ............................................. 70
3.10 FWHM Variability .................................................. ............................................. 71
3.11 Variability of PSF Model Parameters I ........................................................................ 72
3.12 Variability of PSF Model Parameters II ................................................................ 72

4.1 Flow Diagram for a Synthetic Data Generator ................................................................ 83
4.2 Examples of Superellipsoids .......................................................................................... 85
4.3 Filtering Effects on Variance ......................................................................................... 92
4.4 Simulated Renal Data .................................................................................................... 94
4.5 Pre-contrast Signal Intensity Distribution on Kidney Tissue ........................................ 95
4.6 Trans-axial Views of Simulated PET Phantom Data .................................................. 96

5.1 Renal PV Correction Methodology ............................................................................ 102
5.2 High Resolution Tissue Templates ............................................................................. 104
5.3 Arterial ROI and Arterial Input Function ..................................................................... 107
5.4 Tissue Contributions to a Cortical ROI ..................................................................... 108
5.5 Cortical ROI Curves and Estimated Renal Component Curves .................................. 109
5.6 Extraction of Tissue Components from a Large Abdominal ROI ............................... 110
5.7 Robustness of Component Extraction ........................................................................ 111
5.8 Patlak Analysis of Raw and PV Corrected Data ......................................................... 112
5.9 Patlak Results for a Small Cortical ROI ...................................................................... 116

6.1 Effect on a Finite-width Step of its Convolution with a Gaussian PSF (I) .................. 121
6.2 Effect on a Finite-width Step of its Convolution with a Gaussian PSF (II) ............... 122
6.3 Recovery Coefficient for a Cylinder ............................................................................. 124
6.4 Examples of $p(\alpha)$ for Various Recovery Coefficients ........................................ 126
6.5 Inverse Gaussian Fit to Cylindrical Data .................................................................... 128
6.6 Comparison of Several Model-fits to Small Cylindrical Data .................................... 129
6.7 Classification Error vs. CNR (I) .................................................................................. 130
6.8 Classification Error vs. CNR (II) ................................................................................ 131
6.9 Recovery Coefficient for non-Gaussian PSFs ............................................................. 134
6.10 Classification Error vs. CNR for non-Gaussian PSFs ................................................. 135

A.1 Cross-section Diagram of the Renal Parenchyma. .................................................. 144
List of Figures

B.1 Osmotic Gradient in the Medulla ............................................ 149
C.1 Spinning Nuclei, Magnetic Fields, and Precession .................... 153
C.2 Longitudinal and Transversal Relaxation ............................... 154
D.1 Typical PDF and CDF .................................................... 161
Chapter 1

Introduction

The work presented in this thesis is concerned with the development of image processing techniques for improving the diagnosis of renal disease through the use of biomedical imaging methods, in particular, Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI).

DCE-MRI is an upcoming imaging technique that, in the area of renal disease, has the possibility of replacing conventional radioisotope-based imaging, whilst also providing anatomical information alongside the dynamic functional data. This is particularly attractive within paediatrics where, medically, it is highly desirable to reduce the radiation burden of any imaging examination. The ultimate goal of DCE-MRI is to produce absolute image quantification of, in this case, renal function. This thesis presents work that contributes to this overall aim.

Before describing the imaging aspects of this work, an overview of the kidneys and renal function is first presented in this chapter. This is followed by a brief introduction to the most commonly used measures of renal function, and the non image-based techniques used for obtaining them. Whilst some of these techniques can provide robust and highly accurate measures of renal function, image-based methods offer the potential to investigate a wider variety of renal function-related aspects. However, biomedical image data, and in particular MRI data, present substantial image processing challenges for accurate and robust quantitative assessment of renal function. Therefore, having
presented the background and rationale behind the proposed work, the last sections of
the chapter introduce the aims, objectives, and contributions of this dissertation.

1.1 The Kidneys

The kidneys are two small bean-shaped organs that are part of the urinary system. They are located in the posterior abdominal wall (just below the ribcage on either side of the spine), with the right kidney being displaced, slightly lower than the left, by the liver. Each kidney is approximately 11 cm long, 3 cm thick, and 6 cm wide and its weight can range between 135-150 grams in an adult male (Bullock et al., 1994).

The kidneys are primarily responsible for maintaining a constant internal environment despite changes in the external environment, a process known as homoeostasis. This involves maintenance of fluid (i.e. water), electrolyte (i.e. salts and minerals such as sodium and potassium) and acid-base balance, and elimination of waste products from the blood flow. The kidneys accomplish these functions with the production of urine: a solution of waste, toxins, and unnecessary bodily substances that are eliminated from the blood flow. Urine flows from each kidney via a hollow muscular tube, known as the ureter, into the bladder. Urine is then stored in the bladder until excreted from the body via the urethra.

In addition, the kidneys are also important in the production of hormones, such as erythropoietin (which increases the production of blood red cells), and enzymes such as renin (involved in the production of vaso-constricting hormones).

1.1.1 Macroscopic Kidney Anatomy

A diagram of the macroscopic anatomy of a kidney can be seen in Figure 1.1. The diagram shows a cross section of a kidney in which the most significant anatomical features are labelled. A layer of fatty tissue, the perirenal fat, and a thin membrane, known as the renal capsule, envelops and protects the kidney. The kidney itself is composed of two separate structures: the renal parenchyma (i.e. its functional part)
and the renal drainage system. Blood filtration and other processes of urine production, as with most other renal metabolic processes, occur in the parenchyma.

The outer parenchymal region is known as the cortex and is both structurally and functionally different to the medulla (a darker inner region in the diagram). The structural arrangement of the basic functional units of the kidney, the nephrons, is different in these two regions, as they are involved in separate stages of the urine production process; this is further explained in Section 1.2.1. The medullary rays are regions of medullary tissue corresponding to those nephrons located in the outermost cortex. Medullary tissue is arranged in structures known as pyramids, that project towards the drainage system (see diagram) and are surrounded by cortical tissue (cortical arches at the base of pyramids join the cortical columns separating them). Each renal pyramid and its surrounding cortical tissue are known as a renal lobe. From the pyramids urine flows into the renal drainage system through the renal papillae (semi-permeable membranes located at the tip of each pyramid).

The drainage system of the kidney is a hollow tree-like structure where a number of small branches, called minor calyces, collect urine and join to form two or three major
Chapter 1. Introduction

calyces; the mayor calyces open into the renal pelvis, a basin-like structure where all urine is finally collected, which is connected to the bladder through the ureter.

Figure 1.1 also shows the renal artery and vein. These are dedicated vascular vessels that supply/drain blood to/from the kidney through a network of interlobar arteries and veins. The slit-like region through which the renal artery and vein, the ureter, and renal nerves all enter or exit the kidney, is known as the hilum.

In between blood delivery to the nephrons via the renal artery and the collection of urine by the drainage system, there is a highly complex process of urine production, performed by the nephrons in the parenchymal region. A description of urine production in kidney, from a microscopic view, is given in Appendix A through a brief introduction to the nephrons and the vascular network that make up the the parenchyma.

1.2 Renal Function

Urine formation is a combined process of indiscriminate filtration of blood in the renal cortex followed by a more complex processes of selective re-absorption and secretion of water and small molecules in the medulla. Filtration estimation is the most important functional test in renal medicine (Levey et al., 2003), perhaps as it is the easiest to test. Therefore, an overview of renal filtration is presented below. For completeness, the corresponding overview of renal re-absorption and secretion is given in Appendix B. At this point, it is important to note that the functional data obtained through biomedical imaging studies are influenced by urine concentration changes due to re-absorption and secretion and therefore it is important that these processes are understood for any (including filtration) functional image analysis of the kidneys.

1.2.1 Filtration

Plasma filtration occurs in the renal cortex; however, not all parts of the the cortex are involved in this process. The cortex can be divided into cortical labyrinths and the medullary rays. Labyrinths have a convoluted appearance since they are composed of
renal corpuscles, convoluted (proximal and distal) tubules, and their supply vasculature, whereas the medullary rays present a striped appearance from straight tubule segments (such as proximal straight tubules and thin descending limbs) and collecting ducts from the outermost nephrons (see Appendix A for an explanation of the different parts of the nephron). The structural difference between these two cortical regions can be clearly appreciated in Figure 1.2a.

![Microscopic view of cortical tissue](image1.jpg) ![The glomerulus](image2.jpg)

**Figure 1.2:** (a) Differentiated cortical structures: cortical labyrinths, composed of renal corpuscles and convoluted tubules can be easily distinguished from the elongated medullary rays. (b) Schematic diagram of the glomerulus. From Cégep de Rimouski (2007).

It is within the renal corpuscles that blood plasma filtration occurs. Blood arrives to the nephron via the afferent arteriole which branches into a convoluted capillary tuft, the glomerulus, and leaves via the efferent arteriole. Enveloping the glomerulus is the Bowman’s capsule, a double-wall cup-shaped extension of the proximal tubule. Filtration takes place from the glomerulus into Bowman’s capsule (see Figure 1.2b). The filtration barrier is permeable for water, small molecules and ions, but impermeable for blood cells and other large molecules (above 70 kDa) (Peters and Myers, 1998). This barrier is also impermeable for electrically charged proteins. The rate at which plasma filtration occurs in the glomeruli is known as Glomerular Filtration Rate (GFR). Only 20% of the plasma is filtrated as it passes through the glomeruli, with a normal GFR in adult males being of approximately 120ml/min (Gabriel, 1988). This amounts to 170 litres of plasma per day.
GFR is dependent on hydrostatic blood pressure in the glomerular capillaries and it is opposed by hydrostatic ultra filtrate pressure from within Bowman’s capsule. It is also opposed by colloid osmotic pressure in the blood since the membrane is impermeable to proteins.

The kidneys are capable of maintaining a remarkably constant blood flow to the glomeruli (and thus a constant GFR) despite a wide range of systemic pressures. This is very important, since the processes of secretion and re-absorption rely on a steady GFR. If it was too fast, electrolytes would not be sufficiently re-absorbed from the urine within the tubules. If it was too slow, waste substances will be re-absorbed back to the blood flow instead of excreted in urine. The processes by which the kidneys achieve a constant GFR are known as autoregulation. Figure 1.3 shows the range of systemic blood pressures over which the kidneys achieve autoregulation.

**Autoregulation of Glomerular Filtration Rate**

There are several mechanisms that work together in a complimentary manner allowing the kidneys to maintain a stable GFR:

**Myogenic autoregulation** is accomplished through arteriolar muscles. In the presence of an increase of systemic blood pressure, these muscles constrict the afferent arteriole. Conversely, in the presence of a decrease of systemic blood pressure, the arteriolar muscles dilate the afferent arteriole. Thus, this process maintains a constant blood flow to the glomerulus under systemic pressure changes.

**Tubuloglomerular feedback** is a local autoregulation mechanism accomplished via the *juxtaglomerular apparatus*. The juxtaglomerular apparatus is a structure composed of specialised cells found both in the distal tubule and in the glomerulus, as tubules return to the cortex and come into contact with glomeruli. The cells of the distal tubule, the *macula densa* are able to assess the rate of ultrafiltrate flow

---

1 Pressure produced in blood plasma by the proteins it carries. Colloids are plasma proteins and other macromolecules that, because of their large size, are unable to easily cross through capillary walls. This produces an osmotic pressure inside the capillaries that balances out the tendency for fluid to leak out of the capillaries.
by measuring the concentration of sodium ions, $[\text{Na}^+]$, within it. The cells within the glomeruli, the juxtaglomerular cells then act on signals from the macula densa to constrict/dilate the afferent arterioles, thus altering GFR.

In addition to this local effect on renal blood pressure, the tubuloglomerular feedback also has an additional systemic effect on blood pressure since its causes the juxtaglomerular cells to secrete renin, an enzyme involved in the production of angiotensin II, a potent systemic vasoconstrictor which increases GFR.

![Figure 1.3: Autoregulation of GFR. From Imholtz (2008)](image)

Having presented an introduction to the kidneys and the process of renal filtration, the next section provides an overview of the techniques used currently for renal function assessment (using non-imaging techniques).

### 1.3 Non Image-based Renal Function Assessment

There are numerous techniques developed to evaluate renal function. Some of the most widely used non-invasive techniques are based either on the concept of plasma clearance or on the measurement of plasma concentration alone for the estimation of the most important parameters for the diagnosis of renal disease: namely, renal blood flow and glomerular filtration rate.
1.3.1 Measuring Renal Plasma Flow and Renal Blood Flow

The exact Renal Plasma Flow (RPF) rate can be determined using the Fick Principle (after Fick, 1870), using a test substance $X$ that is not stored or metabolised by the kidney:

\[
RPF = \frac{U_{[X]} F_u}{A_{[X]} - V_{[X]}}
\]

(1.1)

where $U_{[X]}$, $A_{[X]}$, and $V_{[X]}$ indicate concentration\(^2\) of substance $X$ in urine, arterial and venous blood respectively and $F_u$ is the urine flow rate (ml/min).

Thus, the clearance of substance $X$ from the blood can be used to calculate renal plasma flow. In adults, RPF is about 650 ml/min (Gabriel, 1985). Sampling arterial plasma to determine concentration is difficult. However, if substance $X$ is chosen so that it is cleared within a single pass through the kidneys, then $V_{[X]}$ becomes zero, simplifying the above equation. Such a substance is para-aminohippurate (PAH). PAH is an organic acid that is both filtered at the glomerulus and extensively secreted by the proximal tubule, resulting in an extraction fraction (i.e. first-pass clearance) that approaches 90%. Therefore, an approximate RPF (underestimated by $\sim$10%) can be found using PAH, known as Effective Renal Plasma Flow (ERPF). Thus:

\[
ERPF = \frac{U_{[PAH]} F_u}{P_{[PAH]}}
\]

(1.2)

where $U_{[PAH]}$ and $P_{[PAH]}$ indicate PAH concentration in urine and systemic/renal arterial plasma respectively. Since PAH is not metabolised and no organ, except the kidney, extracts PAH from the blood, the PAH concentration in the renal artery is equal to the concentration in any non-renal peripheral vein. This can be easily sampled.

Blood is composed of both plasma and hematocrit (Hct), the fraction of blood volume that is composed of cells. Thus, knowing RPF and Hct, Renal Blood Flow (RBF) can be found so that:

\(^{2}\)Concentration is most commonly expressed as Molarity ($M$) = No. of moles/litres of solvent.

Another commonly used measure of concentration is Molality ($m$) = No. of moles/kg. of solvent.
1.3. Non Image-based Renal Function Assessment

\[ RBF = RPF + (RBF \times \text{Hct}) \]

rearranging:

\[ RBF = \frac{RPF}{(1 - \text{Hct})} \]  

(1.3)

ERPF measurements using PAH need a constant intravenous infusion to achieve a constant \( P_{[\text{PAH}]} \) and timed urine samples to measure urine flow. Urine sampling is particularly prone to error, and the whole procedure is time consuming and expensive. Thus, RBF estimation is not widely used in the clinical diagnosis of renal disease. However, accurate measurement of RBF is an active area of research, as this is the most important parameter for the assessment of pathology in the renal vasculature.

1.3.2 Measurement of Glomerular Filtration Rate

As already discussed, GFR estimation is the most important functional test in renal medicine. GFR, the total plasma volume filtered by the glomeruli per unit time, can be measured by the use of some particular substance \( X \) that is freely filtered but is not re-absorbed, secreted, stored or metabolised, so that the amount excreted equals the amount filtered indiscriminately at the glomerulus. Such a substance is inulin, a fructose polymer that is not metabolised and is cleared only by glomerular filtration. Thus, GFR can be calculated using again the concept of clearance such that:

\[ \text{GFR} = \frac{U_{[\text{inu}]} F_u}{P_{[\text{inu}]]}} \]  

(1.4)

where \( U_{[\text{inu}]} \) and \( P_{[\text{inu}]} \) indicate inulin concentration in urine and plasma respectively.

Just as with RBF in the previous section, GFR measurement using the above formula for inulin requires a constant intravenous infusion and has urine collection requirements that are not usually met by patients with compromised renal function. Thus, even though inulin clearance is considered to be the gold standard in GFR measurement, in practice it is only used as a research tool.
Clinically, creatinine is used instead of inulin to determine renal clearance. Creatinine, a product of muscle metabolism, is freely filtered at the glomerulus and is not reabsorbed. However, it is secreted by the tubules and, therefore, the clearance formula overestimates GFR. However, GFR measurement using creatinine clearance does not require intravenous infusion and the clearance period can be extended (usually 24 hours) so that adequate amounts of urine can be obtained. This makes it preferable to inulin clearance for routine medical practice.

Measurement of Glomerular Filtration Rate using Plasma Concentration

To avoid the problems associated with urine collection, GFR can also be estimated from the measurement of the plasma creatinine alone since GFR is approximately inversely proportional to plasma creatinine concentration, \( P_{[cr]} \). The Cockroft-Gault formula (Cockcroft and Gault, 1976) takes into account patient weight and gender to produce an estimate of creatinine clearance (\( C_{cr} \)):

\[
C_{cr} = \frac{(140 - \text{age}) \times \text{weight} \times C_g}{P_{[cr]}}
\]  

(1.5)

where \( C_g \) is 1 for men and 0.85 for women.

There are several variations and modifications on the Cockroft-Gault formula, including other factors, such as the Modification of Diet in Renal Disease (MDRD) formula (Levey et al., 1999) which also accounts for racial origin. Such methods, based on collecting a series of plasma samples and using plasma creatinine concentration formulae, are the basis for GFR measurement in clinical practice (Baker, 1998).

In addition to the plasma concentration methods for estimating renal function using creatinine described above, there are others based on the use of radiopharmaceuticals (radionuclide\(^3\)-labelled substances). Radionuclides were first used to assess renal function by Oeser and Billion (1952), who measured the remaining traces in urine of a

\(^3\)An unstable nuclide that decays either by spontaneous fission or emission of an \( \alpha \)-particle, \( \beta \)-particle, or \( \gamma \)-photon. For medical applications those that decay by the latter process are the norm.
radioactive material that had been administered to the patient. Currently, chromium-51-ethylenediaminetetraacetate $^{51}$Cr-EDTA is used for GFR estimation in clinical practice when higher accuracy than that obtained by creatinine clearance is required since this radiopharmaceutical is not secreted by the tubules. A single shot of the radiopharmaceutical is intravenously injected and plasma is sampled repeatedly, using the same process as in the plasma creatinine method.

The plasma clearance and plasma concentration methods presented in this section yield only global function estimates and very often the relative contribution of each kidney to the total renal function is more important than a single (i.e. combined left and right kidney) global function estimate. Global GFR measures can be deceiving, since there are normally two kidneys in the human body, and one tends to over-compensate for any decreased function that may occur in the other. To overcome this limitation, imaging techniques are commonly used in renal diagnosis.

### 1.4 Biomedical Imaging and Magnetic Resonance Renography

Biomedical imaging has become an increasingly important tool in medical diagnosis from the latter part of the 20th century. Qualitative information from images is routinely used to help in the assessment of a patient’s condition, often allowing the examination of organs and tissues which otherwise would only be accessible through biopsies or surgical intervention. A large emphasis on the research in the field is currently being placed on developing mathematical analysis, modelling, and image processing techniques to extract meaningful quantitative information from these biomedical images that would further improve clinical diagnosis.

A common way to broadly categorising biomedical imaging modalities relates to the type of information acquired: anatomical or functional data. It is possible to acquire highly detailed morphological images through X-ray-based Computed Tomography (CT) and nuclear Magnetic Resonance (MR)-based techniques. Functional information has conventionally been captured through the use of Nuclear Imaging (NI)
techniques that display the spatial distribution of radio-isotope labelled substances through time, such as Positron Emission Tomography (PET), Single Photon Emission Computed Tomography (SPECT), and scintigraphy. However, technological advances are constantly reducing acquisition times in CT and MRI with which it is now possible to obtain functional images that also convey highly detailed anatomical information. Whereas some of these functional methodologies have been adapted from PET and SPECT techniques, there have also been new approaches that make use of phenomena specific to CT or MR, opening new fields of medical and scientific research in functional imaging. For example, the Blood Oxygenation Level Dependent (BOLD) sequences in MRI, that can localise areas of high activity in the brain based on the different MR response from oxygen-laden hemoglobin and oxygen-starved desoxi-hemoglobin, widely used in neuro-science studies.

1.4.1 Renal Imaging

A wide variety of imaging modalities are commonly used in clinical diagnosis of renal disease. A commonly used technique to provide structural information is the Kidney, Ureter and Bladder (KUB) X-ray, used in the detection of renal calcifications and other urinary tract obstructions. Ultrasound techniques are also widely used in detecting kidney stones, renal malformations and urinary tract infections. Another widely used structural imaging technique is the IntraVenous Urogram (IVU), based on the administration of an X-ray opaque substance, can provide two-dimensional (2D) structural information which can be used to detect shape, size or dilation abnormalities, renal cysts, stones and calcifications. The approach can be used to obtain three-dimensional (3D) data using CT techniques.

In addition to anatomical investigations, functional imaging studies are also of great importance in renal diagnosis. Non image-based plasma sampling and/or urine clearance tests produce only a global measure of the combined function of the kidneys. However, biomedical imaging techniques allow the investigation of individual kidney function, either differential (i.e. as the percentage of the total function for each kidney) or single (absolute measurements for each kidney) renal function. To this purpose,
1.4. Biomedical Imaging and Magnetic Resonance Renography

scintigraphy is commonly used to estimate GFR, the most important marker of renal function. The methodologies used for renal function diagnosis by analysing the passage of radiopharmaceuticals through the kidneys using scintigraphy are commonly referred as Radionuclide Renography (RR), and further described in the next chapter.

There are other experimental techniques that are not used routinely in clinical practice. For example, fast CT scanning techniques, such as Electron Beam Computed Tomography (EBCT), have also been used for measuring single-kidney RBF, GFR, and tubular function (Krier et al., 2001). Similarly, PET techniques that allow repeated measurements at short intervals have been used in experimental studies to estimate RBF (Julliard et al., 2007; Ahn et al., 2000). In both NI and CT, patients are subjected to varying doses of ionising radiation.

MRI is also an important technique in clinical diagnosis of renal disease as it is widely used to assess the renal vasculature through MR Angiography studies (Schoenberg et al., 2006). MRI has also been suggested for assessing the urinary tract through MR Urography studies, but perhaps one of the most exiting areas offering great potential is the development of MR Renography techniques to measure GFR, using DCE-MRI. These offer the advantage over NI-based techniques of increased spatial resolution and a lack of ionising radiation, specially desirable in paediatrics and for repeated diagnostic assessment.

Whilst DCE-MRI renography techniques are yet to achieve the robustness and acceptance of NI renography, the potential advantages are appealing as a higher spatial resolution may be used to assess more complex aspects of renal function than GFR. It is the possibility of using MRI data in separately assessing localised kidney function, such as the functional behaviour of the medulla or the cortex, that is driving research into MR renography techniques with the aim to develop more comprehensive methods of measuring kidney function to improve on the current diagnosis of renal disease.

In order to develop effective quantitative analysis techniques for MR Renography, there is a need to understand the image acquisition and storage processes and the effects they produce in the data, such as the relationship between signal intensity and contrast agent concentration, or the effects of digitising and storing a continuous signal in a discrete
grid to form images/volumes. Therefore, the focus of this thesis is to investigate the image processing techniques needed for what is, essentially, a data reduction exercise that attempts to extract meaningful parameters that may be used in a clinical setting (i.e. relative or absolute filtration rate) from large 4D spatio-temporal datasets.

1.4.2 Image Processing for MR Renography

As stated in the previous section, the current technology of choice in clinical studies to provide quantitative evaluation of renal function is NI. Radionuclide Renography is based on tagging a biological compound with a radionuclide which is only extracted from the body through the kidneys. The emitted $\gamma$-rays are captured with a gamma camera to create a set of images that follow the distribution of the compound as a function of time. These data are then used to produce renograms, which are time-concentration curves from manually-drawn renal Regions Of Interest (ROIs). Thus, by comparing left and right kidney renograms, usually through some sort of compartmental analysis, the contribution from each kidney to a total marker of renal function, such as GFR, may be estimated.

MR Renography techniques have been adapted from radioisotope renography, following the same approach, but adapted to address some particular issues of MRI, including:

Movement Correction: widely ignored in RR due to the poorer NI spatial resolution, movement artefacts due to patient breathing are clearly visible and therefore movement correction is usually the first step taken in processing DCE-MRI data.

Segmentation: renal ROIs are drawn on the DCE-MRI data for the extraction of renograms. Depending on the approach, these may be drawn manually or automatically, over 2D images or 3D volumes. Further, these could be drawn either over the whole kidney, the renal parenchyma (i.e. the functional part of the kidney), the renal cortex, or even small subsets of the cortex.

Analysis of concentration curves: the final stage involves the analysis of renogram data, understanding the relationship between MR signal intensities and tracer concentration. There are a wide variety of models with different complexity that
1.4. Biomedical Imaging and Magnetic Resonance Renography

have been proposed in the literature for modelling tracer passage through the kidneys. In general, it is also necessary to measure the renal input function (Buckley et al., 2006) for quantitative analysis, commonly obtained through ROIs placed in the aorta.

The typical segmentation approaches used in MR renography result in binary decisions at every voxel for its inclusion in the ROI. Thus, these approaches assume that the MR signal in every voxel within the segmented region originates from a single tissue. However, the Partial Volume (PV) effect, which arises from the limited intrinsic spatial resolution of an imaging system, results in a mixing of different adjacent intensity classes (often corresponding to tissue types) within a voxel. PV correction methodologies for un-mixing these intensities have been traditionally a concern of MRI (and other modalities such as PET) imaging of the brain, often based in either probabilistic or template-based approaches to extract the different signal component classes. However, the relevance of the PV effect to DCE-MRI data of the kidneys has been previously assumed negligible or routinely ignored.

An important characteristic of biomedical imaging is that there is often a lack of suitable data to validate image processing techniques. For example, automated tissue segmentation results are often compared to manually-drawn contours from clinical experts. These expert-defined data are often referred to as Ground Truth (GT), a term derived from remote sensing applications where image data is verified against real features and materials on the ground, and are assumed to correspond to the true tissue boundaries. However, there is often significant inter and intra-expert variation in manually-drawn GT from clinical data, and it is necessary to make use of non-clinical data, for which more accurate GT can be produced, to quantitatively evaluate methodologies. Typically, such data may be generated from phantom studies or produced through computer simulations.

The areas described above therefore provide the motivation and focus for the work in this thesis. This is explicitly detailed in the next section.
1.5 Aim and Objectives

Given the shortcomings of current imaging methodology and analysis described in Section 1.4.2, the aim of this research work is to improve the current image processing and analysis techniques used in MR Renography in order to provide more accurate and robust estimates of renal function. Taking as an exemplar the problem of GFR estimation, this aim is addressed by the following objectives:

- To quantify the magnitude of the PV effect on a typical DCE-MRI sequence in the kidneys, in particular those of non-renal tissues contributing to renal ROIs.
- To develop a method for correcting the PV effect of non-renal tissues in the kidneys.
- To develop a method for correcting the PV effect in the vasculature.

1.6 Contribution

The contributions of this thesis to the subject can be summarised as:

- **Modelling the effects of acquisition process for a DCE-MRI sequence.**
  The 3D Point Spread Function (PSF) of the acquisition sequence was characterised through phantom experiments at multiple resolutions. This is described in Chapter 3 and published in Rodriguez Gutierrez et al. (2007).

- **Developing a Synthetic data simulator for biomedical data.**
  The simulator was specifically designed to generate, from an initial high resolution dataset, low resolution data affected by PV effects, movement, and tracer passage through different compartments.

- **A method to correct the PV effect on small structures.**
  Arterial vessels, small with respect to the PSF, may be severely affected by the PV effect. A novel method that models and corrects these effects within a probabilistic classification framework is presented in Chapter 6 and published in Rodriguez Gutierrez et al. (2006b) and Rodriguez Gutierrez et al. (2006a).
1.7 Thesis Overview

- A method to quantify and correct the PV effect on the kidneys.
  The PV effect is suggested as an obfuscating phenomena that impedes absolute image quantification. This mixes non-renal components with the true kidney signal. Its effect on renal DCE-MRI data is investigated for the first time and the effects of a partial volume correction stage on renal curves is presented in Chapter 5. Preliminary results, using a 2D PSF model, were presented in Rodriguez-Gutierrez et al. (2006) and a more detailed analysis on 20 kidneys, using a 3D PSF model, can be found in Rodriguez Gutierrez et al. (2008).

1.7 Thesis Overview

This chapter has briefly introduced the kidneys and renal function, followed by an overview of the current non-image-based practices used clinically for renal function assessment: blood and urine sampling-based methods, which can only provide global function estimation. In contrast, image-based methods offer the potential to investigate a wider variety and more detailed aspects of renal function, in particular differential or single kidney GFR.

Chapter 2 presents a detailed introduction image-based renal-function assessment. First, standard renography is described. This NI-based renography technique is the basis of clinical differential GFR estimation, and corresponding MRI renography techniques are largely based on it. The remaining part of the chapter introduces MR renography and the specific issues associated with this technique by a review of the state of the art in the field, including advanced models of kidney function and the image processing techniques required for quantitative renal function estimation: movement correction, segmentation and correction of the PV Effect.

Chapter 3 describes the protocol and acquisition sequences used for dynamic renal imaging. This is followed by work carried out for two preliminary image processing steps to MR Renography. First is the modelling of blurring effects produced by the scanner and imaging sequence on the MRI data, necessary for the PV correction steps introduced later in this dissertation. Second is the implementation of a movement correction technique suitable for 3D DCE-MRI data.
Chapter 4 focuses on the design and implementation of a four-dimensions (4D) synthetic data generator developed for simulating dynamic data and the effects of varying position, tissue noise, and partial volumes. Using a modular approach for these effects, the generator has not only been used to generate synthetic MR data, but also simulated PET phantom data (Rodriguez Gutierrez et al., 2006b).

In Chapter 5 the importance of the PV effect from non-renal contributions on DCE-MRI renograms is investigated. Non-renal contributions from liver, spleen and other surrounding tissues are eliminated from concentration curves derived from typical renal cortical ROIs using a template-based approach. The significance of the the PV effect on filtration estimates, as obtained by analysis of both PV corrected and uncorrected clinical data, is presented.

Chapter 6 considers the PV effect in small structures such as blood vessels. In such cases, the PV effect may be so severe as to blur all pixels within the structure, affecting the concentration curves usually estimated from blood vessel ROIs. A probabilistic partial volume classification technique is proposed, based on analytical modelling of the intrinsic mixing effects produced by the acquisition system on small structures, followed by an intensity recovery step.

Finally, Chapter 7 presents the conclusions drawn from the work presented in this thesis and suggests areas of further work in the field of image processing for MR Renography.
Chapter 2

Image-based Renal Function Estimation

The first part of this chapter provides an introduction to the paradigm of renography. Conventionally, the imaging method of choice has been Radionuclide Renography (RR), based on Nuclear Imaging (NI) techniques, where the passage of an appropriate radiopharmaceutical through the kidneys is followed in time. RR is described in (Section 2.1). The use of Magnetic Resonance (MR)-based renography, as a potential alternative to RR for renal function estimation is then introduced. The overall MR Renography approach has been adapted from the RR methodology, but using Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) instead of NI and therefore using an MR-suitable pharmaceutical that does not emit ionising radiation. Some of the most important considerations for renography specific to the MRI signal are discussed in Section 2.2.

The second half of the chapter introduces the state of the art in image processing and analysis techniques required for quantitative renal function estimation in MR Renography. In particular, these are movement correction (Section 2.3) and segmentation (Sections 2.4 and 2.5) of the DCE-MRI data, and modelling of the pharmaceutical’s passage through the kidney in order to derive meaningful semi-quantitative parameters of renal function from dynamic MR (Section 2.6).
2.1 Radionuclide Renography

Radionuclide Renography techniques are based on the use of a radiopharmaceutical in tracer quantities for the diagnosis of human disease. The standard technique to assess renal function is to use a gamma camera, as developed by Anger in 1958. 2D NI using a gamma camera is also known as scintigraphy, a term historically derived from the use of scintillation detectors (materials that emit visible or near visible light when they absorb ionising radiation) in the gamma camera. The gamma camera comprises a 2D array of scintillator detectors that emit high energy photons when hit by gamma particles. The photons at each detector are then counted, and a 2D image describing the spatial count density is reconstructed. Thus, each image in planar dynamic scintigraphy is a 2D projection of the 3D distribution of the radionuclide in the organs and tissues within the field of view of the detector, with the signal intensity in each pixel being proportional to the photon counts that are, in turn, related to tracer concentration by a linear relationship.

To capture the functional behaviour, a series of sequential images are obtained following the progress of the radiopharmaceutical. The metabolism of the particular agent compound within the kidney can hence be traced from the initial cortical uptake through medullary activity and, finally, urinary excretion.

Renal blood flow, glomerular filtration rate, proximal tubular function, neuromuscular activity (both in the renal pelvis and the ureter), bladder emptying and urinary residual volume can be assessed by separate RR investigations; different radiopharmaceuticals are used depending on the particular aspect of renal function to be investigated. They are composed of a radionuclide and a pharmaceutical with suitable physical and biological properties. The radiopharmaceutical of choice for GFR estimation is Technetium-99m-diethylenetriaminepentaacetic acid ($^{99m}$Tc-DTPA). DTPA was first introduced as a renal agent by Hauser et al. (1970). It has a low extraction fraction (20%) and it is mostly filtrated in the glomeruli. Since it is neither secreted or re-absorbed in the tubules its clearance reflects GFR. Further, the possibility of labelling this compound with $^{99m}$Tc makes it a convenient agent for GFR scintigraphic estimation.

There are many details in the image acquisition protocols in RR such as the radiophar-
2.1. Radionuclide Renography

Figure 2.1: Typical image from $^{99m}$Tc-DTPA radionuclide renography studies (posterior view). Signal intensity is proportional to the number of photon counts, and thus reduced function can be observed in the right kidney when compared to the left kidney, as seen inside their manually drawn ROIs (arrowed). The non-renal ROIs below the kidneys are used for eliminating background contributions from the renal ROIs (see Section 2.1.1). Image from Segami Corporation (2005)

maceuticals used, dosage, timing, administration or otherwise of diuretics, etc. that depend on which aspects of renal function are to be investigated. Nonetheless, these are all variations of the standard renographic technique, known as standard renography, which is briefly introduced below.

2.1.1 Standard Renography

Standard renography allows estimation of both renal clearance and excretion of tracer by the kidneys. Even though there are no recognised methods for the estimation of absolute renal clearance from standard renography alone (Gordon et al., 2001), it can produce estimates of Differential Renal Function (DRF) which can then be combined with blood-sampling global plasma clearance techniques to calculate single kidney absolute function.

A typical NI study last for 20-30 minutes. The radiopharmaceutical is administered
as a bolus\(^1\) to the patient and this is followed by the serial acquisition of images. The standard method to analyse the acquired data is to reduce the 3D sets of data (two spatial and one temporal dimension) by generating time-activity (TA) curves for selected ROIs. These graphical representations of the functional behaviour of the kidney are known as renograms, and are usually displayed along with a sub-sample of the sequential image dataset.

The standard renogram for a healthy kidney differs depending on the radiopharmaceutical used. Nonetheless, a \(^{99m}\)Tc-DTPA renograms presents the following stages: (i) an initial vascular phase, marking the arrival of the bolus to the kidney characterised by a rapidly rising TA curve, (ii) a filtration phase characterised by a lower rise (secretion by the proximal tubules and re-absorption also occur at this stage and this phase will culminate in a peak at which the rate of tracer arrival to the kidney and the excretion rate are equal), (iii) finally, the excretory phase is marked by a decreasing TA curve as no more tracer is delivered to the kidney. A typical \(^{99m}\)Tc-DTPA renogram can be seen in Figure 2.2.

\(1\)A single impulse-like dose of drug administered over a short period of time, often followed by a saline solution.
2.1. Radionuclide Renography

This is usually done through definition of a background ROI in the vicinity of the kidney for which a background ROI curve is generated. Thus, the background-corrected renal curve \( s'_R(t) \) for a renal ROI of size \( A_R \) is given by:

\[
s'_R(t) = s_R(t) - \frac{A_R}{A_B} s_B(t)
\]  

(2.1)

where \( s_B(t) \) is the background curve obtained from an ROI of size \( A_B \), and where \( s_R(t) \) is the uncorrected renal curve.

Once the curves have been background corrected and are thought to represent count rates for the kidneys in time, estimation of renal function can be attempted. There are various methods to estimate DRF. Some of these approaches are described below:

**Integral Method.** This method consists of calculating the Area Under the Curve (AUC) of the background corrected curves. This is simply the cumulative amount of tracer within a renal ROI, \( s_R(t) \):

\[
AUC = \sum_{t=0}^{\tau} s'_R(t)
\]  

(2.2)

The ratio of left and right kidney AUCs is then used as a measure of DRF.

**Deconvolution.** Deconvolution considers the kidney as a Linear Time Invariant (LTI) system. The behaviour of such system is characterised by its impulse response \( h(t) \). The response \( y(t) \) of a LTI system to an input \( x(t) \) is the convolution between the input and the impulse response of the system:

\[
y(t) = x(t) * h(t)
\]  

(2.3)

where \( h(t) \) can be found to be equal to the output of such a system \( y(t) \) when using a unit impulse \( \delta(t) \) as an input:

\[
y(t) = \delta(t) * h(t)
\]  

(2.4)
since the unit impulse can be used, by definition, to conveniently express any function \( f(t) \) as the convolution between itself and the unit impulse:

\[
 f(t) = \delta(t) * f(t) = \int_{-\infty}^{\infty} \delta(t - \tau)f(\tau)d\tau
\]  

(2.5)

where \( \tau \) is just an integration variable.

Modelling the kidney as a LTI system, the output rate \( F_{\text{tracer}}^{\text{OUT}}(t) \) of tracer is related to the input rate \( F_{\text{tracer}}^{\text{IN}}(t) \) of tracer as

\[
 F_{\text{tracer}}^{\text{OUT}}(t) = F_{\text{tracer}}^{\text{IN}}(t) * h(t)
\]  

(2.6)

Thus, if \( F_{\text{tracer}}^{\text{OUT}}(t) \) and \( F_{\text{tracer}}^{\text{IN}}(t) \) are measured, \( h(t) \) can be found by deconvolution of Equation 2.6.

In imaging studies, input rates are derived from cardiac activity curves, and instead of output rates, the quantity of tracer remaining in the kidney, as per background corrected renal ROIs, \( s'_R(t) \), is used (Fleming and Kemp, 1999):

\[
 s'_R(t) = I(t) * H(t)
\]  

(2.7)

where \( H(t) \) is known as the retention function, the fraction of tracer remaining in the organ after a unit impulse. From this function, parameters such as Mean Transit Time (MTT) or GFR can be estimated. Deconvolution of Equation 2.7 is commonly an 'ill-conditioned' problem due to noise, and not easily solved (Sutton and Kempi, 1992; Lawson, 1999). Further, it is important to note the constraint the linearity assumption poses, not for nuclear imaging, but for other imaging modalities such as magnetic resonance. In the case of MR, possible non-linearities in the relationships between signal intensity and tracer concentration (see Appendix C) or flow, may invalidate the model.

Rutland-Patlak This approach was first used by Patlak in 1977 (Gjedde, 1995) to consider the rate of uptake of radionuclide compounds in the brain, but was not published until 1983 (Patlak et al., 1983). Rutland independently used the same
method to measure renal uptake (Rutland, 1979). A two-compartment model is used to model the flow of tracer from the vascular space to the nephrons, with the corresponding tracer concentrations estimated from the signal intensities from 'arterial' and 'renal' ROIs respectively, see Figure 2.3.

The amount of tracer $X$ within a renal ROI, $R_{[X]}(t)$, placed in the parenchyma is assumed to be composed of tracer both in the nephrons $N_{[X]}(t)$ and in the renal vascular space $VR_{[X]}(t)$:

$$R_{[X]}(t) = N_{[X]}(t) + VR_{[X]}(t) \quad (2.8)$$

Assuming unilateral flow from the arterial to the renal space and that no tracer leaves the renal ROI, $N_{[X]}(t)$ may be estimated as proportional to the integral of the tracer concentration curve in an arterial ROI $s_A(t)$, and $VR_{[X]}(t)$ may be estimated as proportional to the amount of tracer from $s_A(t)$. Thus, at a time $t = T$, if $R_{[X]}(t)$ is proportional to the signal intensity in the renal ROI, $s_R(t)$, Equation 2.8 may be re-written as:

$$R_{[X]}(t) \propto s_R(t) = k_1 \times \int_0^T s_A(\tau) \, d\tau + k_2 \times s_A(t) \quad (2.9)$$

where $k_1$ and $k_2$ are constants representing organ uptake and blood background contribution respectively, and $\tau$ is the integration variable. The above formula can be divided by $s_A(t)$, giving the equation for a straight line with gradient $k_1$ and intercept $k_2$:

$$\frac{s_R(t)}{s_A(t)} = k_1 \frac{\int_0^T s_A(\tau) \, d\tau}{s_A(t)} + k_2 \quad (2.10)$$

Using $^{99m}$Tc-DTPA and Rutland-Patlak analysis on the first minute after tracer injection $k_1$ can give an indication of RBF. The same analysis applied between one and two minutes after tracer injection is considered indicative of GFR. After this time, as $^{99m}$Tc-DTPA begins to be excreted from the kidneys, the assumption on which the model is based no longer holds.
Chapter 2. Image-based Renal Function Estimation

Figure 2.3: Rutland-Patlak model: if applied in the filtration phase, $k_1$ represents GFR.

Although many of the methods above give satisfactory DRF estimates under normal global function and symmetrical DRF (40% - 60%), only the Integral and Patlak approaches have been suggested as acceptable for the estimation of differential renal function in cases of poor renal function (Prigent et al., 1999). It can be argued that these renography approaches are limited by the inherent poor data resulting from 2D scintigraphy which requires subjectively-drawn ROIs contaminated from non-specific background activity. In the remaining of this chapter, the potential advantages and disadvantages of MRI data for estimating renal function as well as the state of the art in renography techniques specific to MR are discussed.

2.2 MR Renography

Having looked at nuclear imaging-based renography, this section describes a promising approach, using DCE-MRI techniques which have considerably higher spatial resolution than NI techniques. A brief introduction to the basics of MR can be found in Appendix C. DCE-MRI is based on the effect that paramagnetic contrast agents have in reducing the relaxation time of water protons. In general, this is used to increase the $T_1$ MR signal. As tracer flows through the kidneys a localised increase in signal intensity can thus be observed in dynamic MRI. Hence, MR renograms can be produced as a measure of MR signal intensity (or better, contrast agent concentration) over time for selected ROIs. This section discusses the issues, particular to MRI renography, that need to be taken into account and that impact and ultimately define the type of image processing needed for the acquired MR Renography data.
2.2. MR Renography

The advantage of MR over NI lies in the availability of higher spatial resolution volumetric (rather than projective 2D) data, and thus also eliminating corrupting contributions from under/over-lying background structures. This advantage is used to create new renography methods that characterise localised function of cortical and medullary areas (Fukuda et al., 1996). The diagnostic potential of such techniques is highlighted by von Schulthess et al. (1991) who proposed an imaging sequence and protocol with such high spatial resolution so as to enable estimation of DRF within the different papillae of a kidney. However, for MR Renography to adequately capture renal function processes such as filtration, which occur within a small time window, high temporal sampling (< 2.5 s) is also required. Thus, useful acquisition sequences must provide suitable temporal sampling whilst still producing sufficiently high spatial resolution (de Priester et al., 2003) so as to be clinically advantageous. Figure 2.4 shows typical cortical and medullary renograms from 2D MR Renography.

![Figure 2.4: Cortical (red) and medullary (green) renograms using 2D MR Renography and the ROIs used to produce the curves (bottom-right insert). From 'MRI of the kidney: Glomerular Filtration' presentation by Nicholas Grenier at GOSH, April 2005.](image)

In MR Renography the contrast agent of choice is gadolinium. Just as DTPA is tagged with $^{99m}$Tc in radionuclide renography, the same macromolecule can be tagged with gadolinium to form Gd-DTPA ($C_8H_4GdN_5O_{20}$). Whilst presenting an important advantage in terms of resolution compared to NI, DCE-MRI has the disadvantage of a complex relationship between the MR signal intensity and gadolinium concentration,
[Gd]. This relationship is highly dependent on the MR sequence and tracer dosage. The effects of these two factors on robust and reproducible estimation of renal function, discussed below, is the field that has received most attention to date in the MR Renography literature.

Initial MR Renography studies were limited to qualitative analysis by the $T_2^*$-shortening effect (Kikinis et al., 1987; Fichtner et al., 1994). These studies found that renograms presented a characteristic signal drop during medullary transit, associated with high Gd-DTPA concentrations in the medulla and renal pelvis. The absence of this signal drop indicates reduced ability of the nephrons to concentrate Gd, and thus this qualitative approach was used to differentiate between normal and diseased kidneys. However, quantitative analysis requires that the $T_2^*$ and other effects, which introduce non-linearities to the captured data, are addressed. Some of these are described below.

### 2.2.1 Avoiding the $T_2^*$-shortening Effect in the Kidney

This $T_2^*$-shortening effect is more pronounced in the kidneys than in other parts of the body when using gadolinium. In the kidneys, the filtrate has initially the same concentration as blood plasma (iso-osmotic) but becomes highly concentrated (hyperosmotic) as water is re-absorbed in the renal tubules. Thus, [Gd] is highest in the renal medulla. The typical clinical Gd dose in MRI, 0.1 mmol/kg Gd-DTPA, has been widely found to produce $T_2^*$-shortening effects in the kidney (Grenier et al., 1996; von Schulthess et al., 1991; Kikinis et al., 1987). Therefore, quantitative measurements of renal function may be compromised unless low Gd doses are used. However, lower doses result in a reduced Signal to Noise Ratio (SNR) in MRI, which limits the attainable temporal and spatial resolution. Nonetheless, continuous development of acquisition sequences promises to deliver improved spatial resolution whilst adequately sampling renal function in time.

Quantitative analysis of MR Renography data was first attempted by Taylor et al. (1997), where high correlation between analogous parameters extracted from MR and radioisotope renograms was found. Discrepancies were again attributed to high [Gd] resulting in $T_2^*$-shortening effects. Much research since then, briefly discussed below,
has been directed towards finding an optimal Gd dose that is small enough to avoid $T_2^*$-shortening effect whilst still producing a useful increase on the SNR of the MR signal.

Takeda et al. (1994) performed an in-vitro (phantom) study to find the range at which Gd-DTPA concentration has a linear relationship with signal intensity. They found that signal intensity increase linearly with increased [Gd] up to 2.0 mmol/kg, both for a series of Gradient-Echo (GE) and Spin-Echo (SE) sequences in a 1.5 T magnet (the typical field strength for magnets in clinical applications). In order to avoid the $T_2^*$-shortening effect a reduced dose of 0.05 mmol/kg Gd-DTPA was suggested. This dose has also been suggested by several other authors (de Priester et al., 2001; Giele et al., 2002; de Priester et al., 2003; Buckley et al., 2006) using GE sequences. However, in-vitro results cannot easily be extrapolated to the kidney: Stanisz and Henkelman (2000) demonstrated that the relaxativity of Gd is dependent on the macromolecular content and viscosity of the solution. Thus, Gd in blood plasma or glomerular filtrate presents a higher relaxativity than in a saline solution and, therefore, attempts to quantify the relationship between [Gd] and signal intensity in a saline solution overestimate the true [Gd]. This could mean that the $T_2^*$-shortening effect may appear in lower [Gd] than those quoted above. Further, the concentration ability of the kidney is variable and depends on factors such as the level of hydration (Murakami et al., 1994), which is not always controlled. These uncertainties have led to lower doses, [0.035-0.01] mmol/Kg, being suggested by many authors (Fukuda et al., 1996; Rusinek et al., 2001; Teh et al., 2003; Lee et al., 2003) with consequent lower SNR.

2.2.2 Flow Effect

Another obfuscating issue in quantitative MR Renography is the effect of flow on the MR signal. This effect is well known, and in fact utilised in MR Angiography to obtain remarkable contrast for the vasculature with respect to tissues. In MR Renography, some of the methods used to estimate GFR, e.g. Patlak and other models described in Section 2.6, are based on the extraction of an Arterial Input Function (AIF) from the aorta. Ivancevic et al. (2003a) found that flow effects resulted in significant over-
estimation of signal intensities in the aorta for $T_1$ GE sequences (widely used in MR Renography), and hence this error would propagate to GFR estimates. Further, he suggested a dynamic calibration method to produce flow-calibration curves depending on the administered dose (Ivancevic et al., 2003b).

Having obtained functional data that represents [Gd], the next steps in MR Renography are the application of image processing techniques for the removal of motion artefacts and the segmentation of appropriate renal ROIs from which to extract the renogram. Finally, as in RR, the last step in MR Renography is concerned with the estimation of suitable functional parameters from the renogram. The current state-of-the-art in these three areas of MR Renography is discussed in the remaining of the chapter.

2.3 Movement Correction

Whilst the poor spatial resolution of NI does not capture significant kidney movement due to patient breathing during acquisition, these effects generally produce a clearly visible artefact in DCE-MRI data. Therefore, movement correction is one of the areas within MR Renography that has received considerable attention in the literature. The effects of kidney movement on image data due to patient breathing can be seen in Figure 2.5.

Kidney movement may result from either patient breathing or involuntary movement as they lay in the scanner for several minutes during a MR Renography study. Breath-hold protocols (von Schulthess et al., 1991) and optimisation of the imaging angle to maximise movement artefacts on a single plane can be used to minimise, but do not eliminate these effects.

Initial MR Renography investigations were performed on kidney transplant patients, as these do not suffer from respiration motion artefacts due to the position, lower in the abdomen, of the transplanted kidney. In such cases, and for highly compliant patients, relatively motion-less datasets can be acquired and thus the movement correction stage can be by-passed (de Priester et al., 2001). Another approach, used by Rohrschneider et al. (2000, 2003) was to define ROIs that avoided the upper and lower poles of
2.3. Movement Correction

Figure 2.5: Movement effects due to patient breathing. Two DCE-MRI time-adjacent image slices acquired with a time difference of 2.5 sec. The bounding boxes, drawn at the same location in both images, fully enclose the kidneys from the left image whilst are clearly miss-placed on the right image.

the kidney in 2D coronal slices. Therefore, ROIs would still be placed within the kidney even after small renal movement. However, this approach obviates the fact that uncorrected ROIs would be placed on possibly functionally-different areas of the kidney due to movement. The applicability of the above approaches is limited and therefore adequate motion correction has to be considered for developing a robust methodology for MR Renography, if it is to be used in a routine clinical setting.

Motion correction might be thought of as a multiple image registration problem, estimating a mapping for each data volume in the series, $V_s(x)$, and an some reference data volume, $V_r(x)$. The task of registration is to determine the parameters of a transformation $T$ which, when applied to the data points, $x$, best aligns both volumes:

$$T : x_{V_s} \rightarrow x_{V_r}$$

where $x \in \Omega_{V_s V_r}$ (i.e. the overlap domain, dependent on both volumes and the transformation. To avoid the effects on kidney movement estimation from adjacent structures within the typical Field of View (FOV) commonly used for MR Renography (covering
the whole abdomen), $\Omega$ is defined as a subset of the FOV covering a single kidney and a small area around it. This is often achieved by masking $V_r$ and $V_s$ with binary or grayscale templates to constrain the search space.

The parameters of $T$ are represented by a $p$-dimensional vector, $\alpha$. In MR Renography, $T$ is often considered as a rigid transformation of the kidney, describing motion artefacts due to patient movement and/or internal movement of kidneys due to patient breathing. Therefore, for 2D MR renography ($p = 3$), the parameters defining $T$ are a rotation angle and a 2D translation vector $\alpha = [\psi; t_x; t_y]$. For 3D MR renography ($p = 6$), the parameters defining $T$, are three rotation angles and a 3D translation vector $\alpha = [\phi; \theta; \psi; t_x; t_y; t_z]$. Thus, the problem of movement correction in MR renography is generally posed as that of finding a parameter vector $\alpha(t)$ for each kidney that describes its movement with respect to a reference position. A review of the different published methods for movement correction in MR Renography is presented below.

### 2.3.1 Manual Correction

Manual correction software, where an expert is allowed to translate and rotate the ROI on each image/volume, has been used for MR Renography (Lee et al., 2003). Such an approach suffers from inter and intra-operator variability, and the accuracy expected from visual assessment, specially in the case of 3D MR Renography, is questionable. Further, it is time consuming, with operator time in the range of 2-3 hours for a 4 minutes 3D dataset. However, manual correction or intervention has been commonly used to correct for isolated errors when the automated approaches detailed below have failed.

### 2.3.2 Hough Transform

The initial work on MR renal movement correction was carried out by Gerig et al. (1991) in the early 90s by applying a modified Hough transform to a temporal sequence of 2D images to correct for in-plane translation and rotation.

In this work, the kidney is modelled as boundary contour composed of $n$ points, $M = (m_1, m_2, ..., m_n)$, manually segmented by an expert from one of the images. Once the
kidney has been defined, then for each image \( I_t(x) \) at time \( t \) in the series, edge detection is performed to obtain an edge map, \( I_t(x) \rightarrow E_t(x) \). The model contour is then scanned over the edge map, incrementing the accumulator, \( A \), at \( A(x - m) \) for every point \( m \) in the contour and every boundary point \( x \) in the edge map. The accumulator will present a peak whose coordinates represent the shift coordinates of the model contour relative to the edge map. This process is repeated for rotated versions of the curve model so that in-plane rotations can also be detected.

This approach was used by von Schulthess et al. (1991) to successfully correct kidney movement in up to 75% of the frames on clinical tests, albeit using special breathing protocols during image acquisition. Assessment of correction accuracy was performed through visual inspection. The remaining frames failed to be corrected due to distorted renal shapes resulting from out-of-plane rotations of the kidney, this limitation being inherent to 2D approaches.

### 2.3.3 Distance Minimisation

A widely used approach for movement correction in 2D MR Renography has been that of minimising the 'distance' between two image slices. This approach, very popular within the image registration field where several distance metrics have been investigated, has been applied by various authors (Gupta et al., 2003; Giele et al., 2001; Denis de Senneville et al., 2006a) in MR Renography using (some form of) the correlation coefficient, \( C \). For two images \( I_r \) and \( I_s \), \( C \) is defined as:

\[
C(t_x, t_y) = \frac{\sum_{x,y \in \Omega_r, \Omega_s} (I_r(x,y) - \overline{I_r}) \times (I_s(x-t_x,y-t_y) - \overline{I_s})}{\sqrt{\sum_{x,y \in \Omega_r, \Omega_s} (I_r(x,y) - \overline{I_r})^2 \sum_{x,y \in \Omega_r, \Omega_s} (I_s(x-t_x,y-t_y) - \overline{I_s})^2}} \tag{2.12}
\]

where \( \overline{I_r} \) and \( \overline{I_s} \) are the mean voxel values of \( I_r \) and \( I_s \) within the specified domain. If \( I_s \) is a displaced version of \( I_r \); \( C \) will present a peak, indicating the displacement \( (t_x, t_y) \) required to minimize the distance between the two images. The approach can be extended to include rotation by comparing the peaks resulting from several rotated versions of \( I_s \) and selecting the maximum.
Another distance measure commonly used in medical image registration that has been used in MR Renography (Song et al., 2005) is Mutual Information (MI). In this approach, voxel grey level intensity values are considered as random variables, and the mutual information, $M(I_r, I_s)$, between two images is defined as:

$$M(I_r, I_s) = H(I_s) - H(I_s | I_r)$$  \hspace{1cm} (2.13)

where $H(I_s)$ is the entropy of image $I_s$ and can be interpreted as a measure of uncertainty of the random variable ($H(I_s) = \sum_i p_i \times \log(p_i)$, where $p_i$ is the probability of grey level intensity $i$). The term $H(I_s | I_r)$ denotes the conditional entropy and is based on the conditional probability $p(b | a)$, that is, the probability of grey level intensity $b$ occurring in $I_s$ given that the corresponding voxel in $I_r$ has grey value $a$. Thus, the conditional entropy represents the amount of uncertainty about $I_s$ when $I_r$ is known. Registration in this approach is therefore assumed to correspond to maximizing mutual information as the optimal alignment of images results in minimal conditional entropy.

**Drift Error**

The above approaches are commonly applied to intensity images which are significantly affected by the passage of tracer. Thus, these methods are applied sequentially, so once movement-corrected, $I_s$ becomes $I_r$ for the next image slice. This kind of movement correction, when used for whole-pixel movement, introduces a potential error in the translation vector of ±0.5 pixels in each direction on every image. As these methods are applied sequentially, errors may accumulate producing a drift on the movement estimates that propagates through the sequence until it ultimately fails. A method of addressing the drift problem is the use of an implicit motion model where a 'history' is acquired from pre-contrast enhancement data. Contrast-enhanced data is then compared to each of the entries in the history (Denis de Senneville et al., 2006a) and selecting $T$ from the 'best' matching entry.
2.3. Movement Correction

2.3.4 Phase Difference

Another approach to movement correction of the kidneys is through Fourier analysis, which is a popular approach within the wider image processing field. Fourier transformed images result in complex data which can be decomposed into ‘magnitude’ and ‘phase’ components. In general, most of the information from the original image is contained within the phase component (Oppenheim and Lim, 1981) and, making use of the spatial properties of the Fourier Transform, these can be used to detect motion.

Consider $I_s(x)$ to be a displaced version of $I_r(x)$:

$$I_s(x) = I_r(x - t_0)$$  \hspace{1cm} (2.14)

where $t_0$ is a 2D translation vector. Let $Y_s(w)$ be the Fourier Transform of $I_s(x)$; then, according to the Fourier Shift property:

$$\mathcal{F}(I_s(x)) = Y_s(w) = Y_r(w) \times \exp(-iwt_0)$$  \hspace{1cm} (2.15)

The normalised cross power spectrum is given by:

$$\frac{Y_s(w) \times Y_r(w)^*}{|Y_s(w) \times Y_r(w)^*|} = \exp(-iwt_0)$$  \hspace{1cm} (2.16)

where * indicates the complex conjugate. The translation between $I_s(x)$ and $I_r(x)$ can then be easily found as the Inverse Fourier Transform of Equation 2.16 is a unit impulse, defined at $t_0$.

Giele et al. (2001) presented a displacement correction algorithm based on the phase difference of 2D image slices. Again based on a rigid kidney template manually segmented by an operator, the resulting whole-kidney ROI is used to create an enlarged mask that limits the effects of other structures on the movement detection processes. 2D translation is estimated between consecutive slices $I_r$ when $I_s$. After applying the mask to both images, a Fourier Transform (FT) is applied. The difference in the phase of the transformed masked images is combined with a flat amplitude spectrum and
an Inverse Fourier Transform (IFT) is performed. The result is a surface containing a peak whose coordinates represent the shift coordinates of the search image relative to the reference image. The method was reported to give excellent results over several datasets, with an accuracy of 90% of shifts correctly estimated to within 1 pixel (1.56x1.56mm). However, the method used for assessing these results was again visual inspection.

The phase difference approach was extended to 3D by Song et al. (2005), who also considered rotation. Considering now $V_s(x)$ to be a translated and rotated version of $V_r(x)$, Equation 2.14 becomes:

$$V_s(x) = V_r(R(\phi, \theta, \psi) \times x + t_0)$$

(2.17)

where $t_0$ is a 3D translation vector and $R(\phi, \theta, \psi)$ is a 3D rotation matrix. The magnitude of the Fourier-transformed Equation 2.17 is given by:

$$|Y_s(w)| = |R(\phi, \theta, \psi) \times Y_r(w)|$$

(2.18)

The rotation matrix $R$ is then estimated by optimising the following energy functional:

$$E = \int \int \int (Y_s(w) - R(\phi, \theta, \psi) \times Y_r(w))^2 \, dw$$

(2.19)

Once rotation is estimated, 3D translation is recovered using the same approach as Giele et al. (2001). A further extension to provide 3D sub-voxel movement correction was suggested in the paper, where assuming a successful whole-voxel correction step, spatial displacement is a plane in the phase component of frequency space which can be estimated through plane fitting. Even though plane fitting to noisy phase data has been reported to be inaccurate (Forooosh et al., 2002), Song et al. reported closer results to both simulated and clinical data (manually registered GT) using this method than using MI registration.
2.3. Movement Correction

2.3.5 Assessing Accuracy

A common problem within medical imaging is that quantitatively assessing the performance of an algorithm when lacking GT or reference data. Within motion correction for MR Renography, a wide range of performance assessments have been used. Visual inspection by skilled operator was used by Giele et al. (2001) on 2D data where the position of a kidney contour was judged and corrected if necessary. This method detects both the occurrence and the magnitude of individual errors in the movement correction process. However, this method becomes unreliable when used on 3D data that also includes rotation estimates both in terms of accuracy and precisions.

An alternative approach has been that of measuring a surrogate parameter as an indicator of successful movement correction. For example, various authors have used a decrease on the variance of time-intensity curves (Sun et al., 2004; Song et al., 2005; Gupta et al., 2000) or a better fit to functional models such as Patlak (Denis de Senneville et al., 2006a) for renal ROIs as an indicator of the effects of movement correction. As such, all these approaches do not explicitly illustrate the performance of a movement correction algorithm but rather the benefits of movement correction in obtaining smoother time-intensity curves.

Another approach is that of using synthetic data. Synthetic data intrinsically provides GT, and therefore is suitable for detailed assessment of the performance of movement correction techniques. However, it is important to acknowledge the limitations of the approach and to be aware of its pitfalls. For example, in creating GT data for assessing its subvoxel approach, Song et al. (2005) uses a single real data volume which is translated and rotated to create a series of known transformations, $T$. Using the same volume, a snapshot of the tracer distribution at time $t$, oversimplifies the problem as the effects of a varying tracer distribution as time passes are not captured in these GT data.
2.4 Kidney Segmentation

Having obtained movement-corrected functional data that represents [Gd], the next step in MR Renography is the definition of a suitable ROI from which to extract the renogram. Essentially, there are two problems associated with segmenting MR renography data:

- what tissues to segment or where to place/how to define a suitable ROI for extracting the renogram.
- how is this segmentation to be achieved.

Deciding what ROI to segment is dependent on the analysis to be performed (see Section 2.6). Some authors extract the renogram from whole-kidney ROIs (Gerig et al., 1991; Giele et al., 2001; Song et al., 2005; Lee et al., 2003), whilst others use parenchyma-only ROIs (Hackstein et al., 2003; Buckley et al., 2006). There are other approaches in which separation of cortical and medullary tissues from the DCE-MRI data is attempted (de Priester et al., 2003; Song et al., 2006; Zöllner et al., 2007; Sun et al., 2002), with the aim of employing compartmental models of localised renal function.

In the same manner that the type of renogram analysis determines what renal regions to segment, it determines the image processing techniques employed in segmenting the data. These have ranged from manual drawing of 2D ROIs by a clinical expert to active contour models that employ both the spatial and temporary characteristics of the data to segment functionally distinct areas within the renal parenchyma. A description of some of the segmentation methods in the MR Renography literature follows, after a discussion of the inherent differences and advantages of the 2D and 3D MR Renography approaches, since these directly affect the objectives of image segmentation.

2D and 3D MR Renography

One of the differences between scintigraphic and MR Renography is that, in the former, the number of counts per ROI represents a projected 3D volume of data. That is,
2.4. Kidney Segmentation

Radioisotope renograms indicate total tracer concentration in the kidney as a function of time.

In 2D MR Renography, however, the 2D slice of data can not provide overall [Gd] estimates for the kidney, only a measure per unit (pixel or slice) of tissue. Thus, there is a need in 2D MR Renography to somehow quantify the volume of functional tissue (i.e. the cortex) and multiply it by the per pixel/slice [Gd] in order to estimate DRF or GFR. A method that has been used for approximating the volume of functional tissue is based on the work of Hegeds and Faarup (1972) who suggested that, for normals, cortical volume may be on average approximately 70% of total kidney volume. Annet et al. (2004) measured kidney volumes in rabbit kidneys after sacrifice and used this approach to estimate single kidney GFR. An alternative method to obtain a total kidney volume estimate would be through the use of high resolution CT or MRI data acquired prior to the functional study. However, even if this cortical fraction value was shown to be accurate for some normal cases, it would be preferable to obtain kidney specific estimates from the individual data rather than using a standard value across the population, in order to account for inter-patient variability. Further, this extrapolation of pixel/slice function utilises the assumption that the sampled ROI data is representative of the overall cortex. This may well prove inaccurate, especially in the cases of localised kidney disease (Pedersen et al., 2004).

A solution to the extrapolation problem, fully exploiting the higher spatial resolution of MRI, is 3D MR Renography. In 3D MR Renography, the entire kidney is captured and therefore global (single) kidney estimates of [Gd] are available. Thus, 3D MR Renography can also produce renograms indicating total tracer concentration in the kidney as a function of time. However, 3D volumetric data requires considerably longer acquisition times that single-slice 2D data. Since there is a maximum acquisition time dictated by the need to adequately sample renal function, 3D MR Renography data generally has a significantly reduced in-plane spatial resolution compared to 2D approaches. This, therefore, limits the movement correction, image segmentation, and tracer modelling that can be applied to 3D MR Renography. As it can be seen, balancing temporal and spatial resolution with an appropriate Gd dose is crucial in MR Renography.
2.4.1 Whole-Kidney/Parenchyma Segmentation of Functional Data

Whole kidney or parenchyma segmentation (i.e. excluding the collecting system) is commonly performed in both 2D and 3D MR Renography as it is needed for either a pre-processing step to some of the movement correction methods presented above (mask creation) (Gerig et al., 1991; Giele et al., 2001; Song et al., 2005; Lee et al., 2003), or for defining parenchymal ROIs on which to apply functional models of the kidney as in the work of Hackstein et al. (2003) and Buckley et al. (2006). Most of these approaches rely on manual segmentation by a clinical expert of a single volume, selected to coincide with the peak of cortical enhancement, both for 2D (Gerig et al., 1991; Giele et al., 2001) and 3D (Song et al., 2005; Lee et al., 2003) approaches.

In general, most automated segmentation techniques in MR Renography attempt to use the functional information available on the data not only to perform whole-kidney segmentation, but intra-kidney segmentation also. Therefore, automated segmentation of the whole kidney is discussed in the next section.

2.4.2 Intra-kidney Segmentation of Functional Data

As filtration occurs in the cortex only and some of the functional models used in renography generally used fail to model correctly the medulla (i.e. Patlak), a popular approach is to attempt to segment the cortex, medulla, and collecting system. Segmentation of low resolution renal functional data is inherently difficult due to blurred edge data, particularly away from the centre-most slices in 3D DCE-MRI.

Some authors have manually segmented a single 2D slice from an appropriate time point in the sequence that maximises the cortico-medullary junction (Annet et al., 2004). A somewhat rudimentary automation approach was proposed by de Priester et al. (2001), who after segmenting the whole kidney through the subtraction of pre-enhancement from post-enhancement mean images performs a series of (mathematical) morphological operations to heuristically define the cortex and the medulla. Some, more refined attempts are described below.
2.4. Kidney Segmentation

Temporal Segmentation of the Kidney Through Clustering

A possible approach, used both in MR and PET, is to use the temporal and spatial information available utilising cluster analysis of signal intensity/concentration curves for each pixel to segment the data set. Song et al. (2006) and Zöllner et al. (2007) used this method to segment the kidney, under the assumption of having an *a priori* known number of possible concentration curves in the kidney: ‘cortex’, medulla’ and ‘collecting system’ (Song et al., 2006; Zöllner et al., 2007).

Another approach, not used in MR Renography, but worth describing here, is Factor Analysis (FA). This technique, commonly used in NI-based functional techniques to separate functionally distinct groups of pixels, is based on the extraction of a series of factors (TA curves) and associated images (coefficients at each voxel location). Thus, each factor normally represents the functional behaviour of a different organ or tissue (kidney, blood pool, etc.). The separation of the different factors is accomplished by applying Singular Value Decomposition (SVD) to a matrix of dixels (a dixel, from ‘dynamic pixel’, is simply a vector, $i_x$, defined for each voxel $x$, consisting of the intensities of that pixel along the time sequence). These type of technique has been used in PET, for example, to determine RBF (and therefore the factor image containing the cortical area involved in this function) (Ahn et al., 2000).

Temporal segmentation of the Kidney Through Active Contours

An approach to kidney segmentation that attempts to use the temporal information available across the whole dataset was presented by Sun et al. (2002) for 2D MR Renography. The method was not only used to segment the kidney, but to segment the cortex from the medulla based on their differentiated temporal dynamics. For each pixel $(x, y)$, they define a dixel, $i_{(x,y)}$. Further, they define $C$ as the boundary curve that divides a subset of dixels from its superset (i.e. cortex from the rest of the kidney), with $\Omega_1$ representing the inside of the curve and with $\Omega_0$ the outside. Segmentation is then performed by minimising the following energy functional:
\[ E(C) = \mu \times L(C) + \lambda_1 \int_{\Omega_i} \text{dis}(i(x,y), \bar{r}_i) \, dx \, dy + \lambda_2 \int_{\Omega_o} \text{dis}(i(x,y), \bar{r}_o) \, dx \, dy \]  \hspace{1cm} (2.20)

where \( \bar{r}_i \) and \( \bar{r}_o \) are the average intensities inside and outside of the curve respectively, \( \text{dis()} \) is a distance metric between two vectors, \( L(C) \) is the length of the curve, and the parameters \( \mu, \lambda_1, \) and \( \lambda_2 \) are positive scalars. The first term penalises the length of the curve, whilst the second and third maximise the distinction between the two regions.

This approach uses both temporal information in the distances within the vectors which sum over the sequence and spatial information by integrating over the image regions. Thus, the proposed energy functional is an interesting take on the popular image segmentation field of active contours (Blake and Isard, 1998). However, the approach was only applied to 2D data of rat kidneys.

2.4.3 Non-rigid Kidneys

Whilst all the approaches above are based on rigid kidneys, a segmentation approach based on a deformable kidney was suggested by Sauce et al. (2006). Starting with a manually segmented renal contour, a 3D warping transformation using B-spline basis functions is estimated based on the optimisation of a MI distance metric. However, there is no evidence in the literature that human kidneys are deformed at all under normal breathing. Even in the event that there were small deformations, there is no evidence that these might be observable at typical DCE-MRI resolution. Arguably, it is more likely that any non-rigid transformations estimated are the result of changes in tracer bio-distribution following tracer passage from the cortex to the medulla than due to kidney deformations.

An important point to make at this stage is that by defining a deformable contour, further processing using pixel-based techniques cannot be used. For example, in the companion work to Sauce et al. (2006), the authors go on to attempt segmentation of the cortex, medulla, and collecting system through clustering of pixel-based intensity curves (Zöllner et al., 2007). This assumes that each voxel represents the same amount
of functioning tissue at each time point, something which is explicitly invalidated by defining a non-rigid kidney.

2.5 Partial Volume Effect

In the previous section, a review of the segmentation techniques used in MR Renography was presented. All the approaches discussed are based on taking a binary decision at voxel level to determine whether the voxel belongs or not to a particular region or tissue class \( \omega_c \). Whilst this may be a reasonable approach to take in non-boundary voxels, grey level intensities near tissue boundaries are likely to result from the combination of the MR signal more than a single tissue class. This mixing of the signal from different adjacent classes is known as the Partial Volume effect. Even though the PV effect has been largely ignored by the MR Renography literature, it is highly likely to have substantial influence for quantitative analysis of renal function. Therefore, an introduction to the PV effect is presented in this section.

The PV effect arises from the limited intrinsic spatial resolution of an imaging system. The finite bandwidth of acquisition systems limits the representation of high frequency information within the image. The information conveyed by high spatial frequencies enables accurate portrayal of edges and small image structures. Therefore imaging systems that fail to adequately capture high frequencies produce blurred representations of the object. Detector sizes, digitisation into a discrete grid, and storage size limitations further contribute to potential corruption of the continuous signal. As a result, grey level intensities of voxels in the boundary between tissues are generally the result of signal mixtures from those tissues. The PV effect is illustrated in Figure 2.6 for the 2D boundary between two regions of different signal intensity. The blurred representation of the edge, when binned into a discrete grid, results in boundary pixels exhibiting false intensities that are clearly not present in the original regions but are the result of a mixture between the original components.

In MR Renography, the time constraints associated with capturing large volumetric datasets covering both kidneys and the aorta within a typically small sampling time result in relatively low spatial resolution where the PV effect is present and likely to
Chapter 2. Image-based Renal Function Estimation

Figure 2.6: The Partial Volume effect: two distinct regions characterised by different signal intensity (left) acquired with a finite bandwidth imaging system result in a blurred representation of the boundary (middle). The signal intensities of the boundary pixels in the resulting discrete image (right) are the result of a mixture of the original intensities.

have substantial influence for quantitative analysis of renal function. The only work to address this issue in the renal MR literature is that of Giele et al. (2002). The authors attempt to find the percentage of cortical tissue on a parenchymal ROI by equating it to the ratio between enhancement slopes of a cortex-only and a parenchymal ROI. Whilst there are no other works in the MR Renography literature that address this issue, the PV effect has been widely addressed both in MRI imaging of the brain and other imaging modalities such as PET. A brief introduction to some of these approaches is presented below.

2.5.1 Template-based PV Classification

A significant body of work addressing the PV effect can be found in the PET and SPECT literature, where, due to the low spatial resolution that these technologies offer, ~ 5 mm Full-Width at Half-Maximum (FWHM), there is consensus for the need of PV effect correction in quantitative analysis involving small structures (Quarantelli et al., 2004).

A number of approaches in this field are based on the definition of high-resolution binary
templates assumed to provide accurate anatomical representations of organs or tissue boundaries. With respect to functional PV correction, these approaches inherently assume that the functional information can be confined to corresponding anatomical regions. These templates are generally obtained from segmentation of high resolution anatomical data from CT or MRI scans. The general approach consists of convolving these high resolution templates with the PSF of the PET/SPECT scanner, down sampling to the same resolution as the PET data, and registering the resultant composite data with the functional PET/SPECT data. The superposition of these convolved templates onto the corresponding functional regions produces boundary overlaps that represent a mixing vector for each voxel, the components of which are assumed to represent the influence of adjacent functional tissue classes. Within a particular tissue, as defined by its high resolution template, voxels that present adjacent contributions are considered no longer pure tissue voxels. The combination of all the exogenous contributions to a voxel is commonly referred to as underspill. Similarly, contributions from that tissue contaminate voxels outside the tissue template boundary, are referred to as overspill. After registration with the PET data, each voxel’s value in a convolved tissue template represents that particular tissue’s contribution to the voxel.

Underspill values, calculated as above, may be used to correct for the FV effects in PET data for a ROI placed on a particular tissue by weighting the signal intensities in the ROI with the corresponding voxels in the convolved tissue template (Meltzer et al., 1990). A generalised approach was proposed by Rousset et al. (1998a) where the signal intensity in each voxel is defined as a weighted sum of individual (pure) tissue intensities. The weights represent each tissue’s contribution to the voxel as defined above, and can be arranged to form a matrix, known as a Geometric Transfer Matrix (GTM). The matrix can then be inverted to solve a system of equations allowing the estimation of the individual tissue intensities contributing to each voxel.

The above methodology have been shown to produce promising results in various studies (Quarantelli et al., 2004; Frouin et al., 2002). Among the sources of error for the approach, miss-registration is often considered to have the greatest effects (Rousset et al., 1998b; Meltzer et al., 1999). Correction for PET data relies on a priori anatomical information from high resolution imaging modalities such as MR and inter-modal
registration has often to cope with different resolutions, alignment, FOV orientations, and patient positioning. When considering such an approach for MR functional data, such as in the case of MR Renography, however, advantage can be taken in acquiring both the anatomical and functional data within a single study, therefore simplifying the registration problem.

### 2.5.2 Probabilistic PV Classification

Another approach for partial volume correction, which does not use a priori anatomical information and has received considerable interest, is that of probabilistic PV classification. Much of the work in this area is based on finite mixture modelling, using Bayes theorem to calculate the posterior probability $P(\omega_c|x)$ of a voxel belonging to a particular tissue $\omega_c$ given a particular voxel measurement $x$:

$$
P(\omega_c|x) = \frac{p(x|\omega_c)P(\omega_c)}{p(x)} \tag{2.21}
$$

where $p(x|\omega_c)$ is the likelihood of measurement $x$ originating from class $\omega_c$. $P(\omega_c)$ is the scalar prior probability of class $\omega_c$ occurring. The denominator $p(x)$ is a normalisation term which is the sum of each class’s likelihood scaled by its prior probability:

$$
p(x) = \sum_{\omega_i} p(x|\omega_i)P(\omega_i) \tag{2.22}
$$

Likelihood models attempt describe the probable range of frequencies for some qualities from the measured signal information that may be used to define statistically distinct image regions (normally associated with different tissues or areas of distinct functional activity). Some of the voxel measurements, $x$, commonly used to describe distinct image regions in medical images range from simple 1D intensity or gradient vectors, to more complex multidimensional feature vector descriptors (Williamson et al., 2002; Chiverton and Wells, 2005, 2004).

A common approach is to model likelihoods for pure tissue classes as Gaussian distributions (Choi et al., 1991) where $\sigma$ describes the signal variability due to tissue heterogeneity, noise, and the acquisition process. Within this framework, PV affected
data is modelled as a separate class with its own PV likelihood model. Another approach is to attempt to model the intrinsic mixing caused by an idealised (noiseless) acquisition system via PV mixing prior densities. These two approaches are discussed below.

**PV-Likelihood Models**

Santago and Gage (1993) modelled a two-class problem with Gaussian distributions, $G(\mu_a, \sigma^2)$ and $G(\mu_b, \sigma^2)$, with equal variance for the pure tissues, and added an extra PV class whose likelihood was the result of convolving the pure tissue distribution with a uniform distribution defined between the means of the pure classes $\mu_a$ and $\mu_b$. The above approach assumes the same variance for both pure tissues. A modification capable of modelling distinct variances for each class is presented by Vokurka et al. (2002) where the uniform PV prior is decomposed into two equal but opposed triangles each convolved with its corresponding Gaussian. Therefore, this approach results in two separate likelihood models for the PV data. Examples of these two different approaches for modelling the PV likelihood distribution can be seen in Figures 2.7 and 2.8.

**Figure 2.7:** PV likelihood models for the distribution of intensities that occur as a result of the PV effect. The two scaled pure distributions (solid lines) and the (scaled) PV distribution (broken line), created by convolving with the pure distribution with a uniform distribution, produce the (normalised, i.e. unscaled) overall PV likelihood (Santago and Gage, 1993) represented by a dotted line.
Figure 2.8: PV likelihood models for the distribution of intensities that occur as a result of the PV effect. The two scaled pure distributions (solid lines) and the (scaled) PV distribution (broken line), created by convolving each pure distribution with a corresponding triangular function, produce the overall PV likelihood (Vokurka et al., 2002) represented by a dotted line.

PV-Mixing Prior Models

By contrast to the above, PV effects may also be modelled through the use of mixing prior densities, $p(\alpha)$. These model the mixing that would be produced in the absence of noise for homogeneous tissues and therefore describe the intrinsic effects that the acquisition system has on noiseless data. Some authors have used uniform distributions (Laidlaw et al., 1998), but this models the PSF as an averaging function, which is an unrealistic model for MR data. A number of researchers have investigated the use of other non-uniform mixing priors to model PV effects. These have ranged from heuristic functions, such as the Beta distribution used by Kitamoto and Takagi (1999) for remote sensing applications and Benford's Law (Wells et al., 2007) (for MR data), to analytically modelling the effects that a particular PSF might produce on an edge between two tissues (Chiverton and Wells, June 2006). A comparison of the mixing functions from these approaches may be seen in Figure 2.9.
Chiverton and Wells (June 2006) developed this model for an imaging system with finite bandwidth and spatially-invariant Gaussian PSF, $G(\sigma)$, such as found in PET data. The effects of convolving this PSF with an idealised noiseless infinite-width step with unit height (simulating the edge between two sufficiently large areas of distinct tissue with respect to the size of the PDF) is given by:

$$g(x) \simeq \frac{1}{2} \left[ \text{erf} \left( \frac{x}{\sqrt{2}\sigma^2} \right) + 1 \right]$$  \hspace{1cm} (2.23)

An estimate of $p(\alpha)$ can be obtained by inverting $g(x)$ and finding its gradient. This is referred to as the Inverse Cumulative Gaussian (ICG) (Chiverton and Wells, June 2006):

$$p_{ICG}(\alpha) = g^{-1}(x) = C_{ICG} \times \sqrt{2\pi\sigma^2} \times \exp(\text{erf}^{-1}(2\alpha - 1)^2).$$  \hspace{1cm} (2.24)

where $C_{ICG}$ is a normalising term for the above model and $\text{erf}^{-1}()$ is the inverse error function (Carlitz, 1963). By providing an analytic framework, the approach of Chiverton and Wells (June 2006) can be extended to other imaging modalities with non-Gaussian PSFs, or other (non-infinite) edge models.

Figure 2.9: Comparison of the three (unscaled) functions used in prior mixing modelling: Beta function (solid line), Benford mixing density (broken line) and the ICG mixing density (dotted line).
Likelihood models are only part of the problem to be considered in probabilistic PV classification. All the above approaches make use of mean and variance estimates for each tissue inferred from image data. In MR, the use of training data from previous datasets is not usually possible due to the inherent variability of the image data from different persons, scanners, and imaging sequences. Therefore, estimation of these parameters from the available data in the image is normally undertaken in order to maximise the overall class posterior probability, i.e.:

$$P(\omega_c|X) = \prod_{\forall x} P(\omega_c|x)$$

(2.25)

This overall probability is maximised by determining optimal parameter estimates from the image data through iterative strategies such as the Expectation-Maximisation algorithm, or through stochastic techniques (van Leemput et al., 2003).

### 2.6 Modelling Kidney Function

Having presented MR Renography and investigated some of the advantages and disadvantages that it poses with respect to RR, followed by an introduction to the image processing techniques required for MR Renography, this last section addresses the analysis of the MR renogram for the estimation of kidney function.

GFR may be calculated from $[\text{Gd}]$ in MR Renography through models of kidney function such as those used in radioisotope renography, adapted for use in MR Renography. The Integral method (see Section 2.1.1) has been used by Teh et al. (2003) for DRF measurements in clinical studies. The same approach was also used by Rohrschneider et al. (2000, 2003) to estimate both DRF and urinary excretion in children. The Rutland-Patlak method has also been used by Pedersen et al. (2004) to measure differential RBF and by Hackstein et al. (2003); Buckley et al. (2006) to measure single kidney GFR. These two studies report high correlation for estimates of single kidney GFR between MR and reference (image and non image-based) techniques.

However, the higher spatial resolution of MR allows for the use of more detailed models, separately characterising cortical and medullary function. Thus, several methods have
been proposed for the analysis of localised renal function which may provide powerful functional diagnosis tools than DRF or single kidney GFR estimation. Some of these are described below.

2.6.1 Multi-compartment Renal Models

The Rutland-Patlak model assumes that no tracer leaves the renal ROI during the sampling period, and generally ROIs that encompass the whole renal parenchyma are selected either to comply with this assumption or because separating the cortex from the medulla is considered too challenging (Buckley et al., 2006). However, the inclusion of medullary components that are not involved in the filtration may introduce substantial errors in the Rutland-Patlak model. Thus, using the high spatial resolution available in MR, Annet et al. (2004) suggested a two compartment model (based on the Rutland-Patlak model) that uses a cortex-only ROI instead of whole-kidney ROI and includes an exponential decay term to model the signal loss associated with tracer leaving the ROI.

However, within cortical-only ROIs there are also areas of tissue that do not participate in filtration, such as the medullary rays. Even within the cortical labyrinths themselves, the convoluted tubules are not involved in filtration. Since corpuscle-only ROIs cannot be determined at the current level of MR spatial resolution, all these components contribute to errors in GFR estimation. To take account of these effects, more complex models attempting to characterise further the process of urine production have been presented. Rusinek et al. (2004) included compartments for the glomerular capillary, proximal convoluted tubules, loops of Henle, distal convoluted tubules, collecting ducts, and the collecting system. A schematic diagram of this model can be seen in Figure 2.10 below.
2.6.2 Localised Tracer Kinetics in the Kidney

By using complex renal compartmental models such as those described above, observed whole-kidney, cortical, or medullary curves can be de-composed into component curves that correspond to the tracer kinetics in each compartment. These sub-component curves may then be used to extract more accurate data on different renal aspects than the overall renal curves (e.g. proximal tubule compartments for GFR, or arterial compartments for RBF).

A possible method for modelling the tracer kinetics in these compartments is to use gamma functions. Gamma functions are commonly used to model flow data as they present a rapid increase rate followed by an exponential decay. Further, they can easily be characterised by only three parameters: the time-to-peak, $\tau$, peak value, $\mu$, and decay rate, $\lambda$. A gamma function, $\gamma(t)$, is defined as:

$$\gamma(t) = \mu \times \left(\frac{t}{\tau}\right)^{\lambda} \times e^{\lambda(t-\tau)} \quad \text{for} \quad t > 0 \quad (2.26)$$

The variation in the type of shapes than can be generated using such a function can be seen in Figure 2.11.

Based on the work of Choyke et al. (1989) and Wolf et al. (1994), a renal tracer kinetics model based on gamma functions has been suggested by de Priester et al. (2003). It assumes that the cortex and medullary concentration curves may be composed of four
2.6. Modelling Kidney Function

Figure 2.11: Gamma variate function: Parameter \( \mu \) is a scale factor for the value of the function, \( \tau \) specifies the peak time, and \( \lambda \) the shape of the curve. Here, \( \mu = 1, \tau = 20, \) and \( \lambda = 0.1 \) (solid line), 1 (dots) and 2 (broken line).

components each, although the medullary components are delayed with respect to the cortical components and that in some cases they refer to different physiological stages than those of the cortex. In the cortex, the components are (a) vascular, (b) proximal tubules, (c) distal tubules, and (d) washout (corresponding to re-circulating tracer). In the medulla, they are: (a) vascular, (b) descending tubules, loop of Henle, and ascending tubules, (c) collecting ducts, and (d) washout.

Taking the cortex model as an example (the medullary one is constructed in the same manner), the first three components \( \gamma_n \) for \( n = 0, 1, 2 \) are modelled using gamma functions:

\[
\gamma_n(t) = \mu \times \left( \frac{t}{\tau_n} \right)^{\lambda_n \tau_n} \times e^{\lambda_n (\tau_n - t)}
\]

where \( \tau_n \) are the time-to-peak delays for the vascular, proximal, and distal compartments respectively. The washout component is modelled by, instead of a gamma function, a slower-rising merging exponential, \( m_e(t) \):

\[
m_e(t) = A \left( 1 - e^{-t/\tau_3} \right)^2 \times e^{-\lambda_3 t}
\]

where \( \tau_3 \) is the time to peak, \( \lambda_3 \) controls the width of the merging exponential modelling
the washout phase, and $A$ is a normalising factor that sets $m_e(T_f) = 1$, $T_f$ being the final time sample in the sequence.

The four curves are combined together to form an enhancement curve $E(t)$, corresponding to $[\text{Gd}]$ as defined in Equation C.3:

$$E(t) = \gamma_0(t - t_a) + \gamma_1(t - t_a - \tau_0) + \gamma_2(t - t_a - \tau_0) + \mu_3 m_e(t - t_a - \tau_2) \quad (2.29)$$

where $t_a$ is the time of arrival of tracer to the kidney. The correction in $\gamma_1$ and $\gamma_2$ for $\tau_0$ is because the vascular phase is assumed to be the input to both the tubular phases.

These type of functional models have also been used across other imaging modalities than MR. For example, the fitting from a similar model (cortical model only, no washout stage) (Juillard et al., 2004) to CT data can be seen in Figure 2.12. Whilst using a different tracer, and therefore producing a different enhancement curve to those expected from Gd, the example shows the convenience of gamma function-based models in representing tracer kinetics data.

![Figure 2.12: Three-component cortical uptake model consisting of vascular, proximal tubule, and distal tubule components fitted to CT functional pig data. From (Juillard et al., 2004).](image)

The problem of fitting these model functions to the data is solved through optimisation techniques that minimise some distance measure between the model and the data,
using domain knowledge of the problem (e.g. the vascular peak precedes the proximal peak, which in terms precedes the distal peak, etc.) to constrain the set of possible solutions. As can be seen, modelling tracer kinetics in such detail offers the possibility of investigating localised aspects of renal function in the renal cortex and medulla.

2.7 Conclusions

In this Chapter, an overview of MR Renography has been presented. Prior published work, alongside ongoing activity in the medical imaging community addresses some of the problems related to inherent limitations of the tracer and MR sequences used ($T_2^*$-shortening effect, conversion of SI to [Gd], flow effects), and to the models that are applied to the data for estimating GFR. With new MR scanners and pulse sequences being developed, these are perhaps the most variable aspects to be taken into account in MR Renography. MR Renography image processing and analysis algorithms should, therefore, be as general as possible in order to accommodate for differences in the acquired MR data.

In the second part of the chapter, the key image processing-related problems that must be addressed for the routine clinical use of quantitative MR Renography have been presented. Firstly, movement correction, required to remove the noise that otherwise would appear on renograms due to both patient breathing and involuntary movement. Secondly, kidney segmentation, required for defining renal ROIs from which to obtain renograms, and for calculating kidney volume or the amount of functional tissue. Thirdly, the Partial Volume effect. This effect has been widely ignored in MR Renography, but is suggested here, for the first time, as a significant obfuscating factor in renal function estimation in MR Renography. Consequently, most of the work in this dissertation is aimed at quantifying and correcting the PV effect in MR renography studies.
Chapter 3

MR Data Acquisition and Preprocessing

This chapter describes the protocol and acquisition sequences used to capture the Magnetic Resonance (MR) data used in this work and preliminary processing concerned with modeling the Point Spread Function (PSF) of the acquisition sequence and movement correction of these data.

Section 3.1 describes the sequences used both for the acquisition of Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) data and anatomical MR data, taken as part of the protocol prior to the contrast enhanced sequence. These anatomical data are used for estimating and correcting the Partial Volume (PV) effect on the DCE-MRI data. Section 3.2 investigates the relationship between Gadolinium concentration and signal intensity for the suggested DCE-MRI data acquisition sequence and Section 3.3 presents the work undertaken on phantom data to model the PSF of the dynamic data acquisition sequence. Finally, Section 3.4 presents a 3D movement correction approach for DCE-MRI data.

3.1 Data Acquisition

There are many acquisition sequences that have been proposed for DCE-MRI of the kidneys in the literature, as in the works described in Chapter 2 and Appendix C.
The approaches are varied and usually tailored to highlighting aspects of a particular research interest. In attempting to quantify GFR, the most important factor in deciding the characteristics of the acquisition sequence of dynamic data is the trade-off to be made between temporal and spatial resolution: it is important to maximise spatial resolution to take advantage of the benefits of MR, though acquisition time should be small enough to ensure appropriate sampling of the renograms and arterial input curves.

The acquisition protocol used for this work is based on the acquisition of 3D abdominal oblique-coronal oriented DCE-MRI data volumes captured every 2.5 s for > 5 minutes. This sampling time was thought to provide a good balance between spatial resolution requirements and an appropriate characterisation of the SI curves, in particular the AIF. The oblique-coronal orientation, along the long axis of the kidney was used to minimise movement artefacts. The 3D data volumes capture an abdominal region that encloses both kidneys completely and also extends posteriorly to include the aorta. In addition to the dynamic scan, an anatomical scan is acquired prior to contrast administration, using the same FOV. All data were acquired from healthy volunteers with normal renal function at Great Ormond Street Hospital on a 1.5 T Siemens Avanto scanner. Both dynamic and anatomic data acquisition sequences are described below.

**Dynamic Data**

DCE-MRI data were acquired using a Spoiled Gradient Recalled (SPGR) 3D Fast Low Angle Shot (FLASH) Volumetric Interpolated Breath hold Examination (VIBE) pulse-sequence (Rofsky et al., 1999): TE/TR = 0.53/1.63 ms, flip angle = 17, acquisition matrix = 128x104 voxels, 400x325 mm² FOV. Thus, the dynamic dataset consisted of 3D volumes with 18 slices of 7.5 mm thickness (no gap) and an in-plane voxel dimension of 3.1x3.1 mm², acquired every 2.5 s for several minutes. The injected Gd-DTPA dose was 0.05 mmol/kg body weight, injected as a bolus at 2 ml/s using an automatic injector (Spectris). The contrast agent bolus was immediately followed by a 15 ml saline flush injected at the same speed. Several image slices of a volunteer at key time points during the scan can be seen in Figure 3.1.
3.2 Linearity of Signal Intensity vs. Gd Concentration

Recently, this sequence has also been used by Sance et al. (2006) for GFR estimation using MR Renography. It is important to note that although adequate for GFR estimation, this type of sequence might not be suitable for investigating other aspects of renal function such as RBF due to insufficient temporal sampling. Thus, MR Renography sequences are usually considered in the context of specific aspects of renal function.

Anatomic Data

The anatomical high-resolution scan was acquired during breath-hold using a true-Fast Imaging with Steady state Precession (FISP) sequence: TE/TR = 1.67/3.34 ms, flip angle = 68, acquisition matrix = 256x187 voxels, 400x325 mm² FOV. Thus, the high-resolution anatomical data consisted of a 3D volume with 18 slices of 7.5mm thickness (no gap) and an in-plane voxel dimension of 1.56x1.56 mm². An exemplar image slice a volunteer can be seen in Figure 3.1.

3.2 Linearity of Signal Intensity vs. Gd Concentration

As highlighted in previous chapters, the relationship between signal intensity and Gd concentration needs to be considered in quantitative MR Renography. An experiment was carried out using the proposed acquisition sequence on a phantom with several dilutions of the Gd contrast agent used at Great Ormond Street Hospital (Magnevist). Figure 3.2 shows the results. The relationship between $T_1$ relaxation and [Gd] remains linear for the range of concentrations used in the experiment [0.1-3.7] mmol/L.

3.3 Determining the PSF of the Dynamic Data Sequence

The process to estimate the blurring produced by the dynamic data acquisition sequence on the MR signal is described in this section. Initially, this 3D blurring is formally defined through the PSF and its 1D approximation, the LSF. The process of accurately estimating LSFs from noisy data is then described. A 1D model for the LSF is empirically derived from experimental phantom data captured using the proposed
3.3.1 Point Spread Function

A linear, shift-invariant imaging system can be completely characterised by its PSF, i.e. the system response to a unit impulse. Since any 3D signal, \( f(x, y, z) \), may be described as a series of shifted infinitesimal impulses, the output, \( o(x, y, z) \), of a linear, shift-invariant system can be considered a sum of PSFs, \( h(x, y, z) \), each shifted and scaled according to the location and height of the corresponding impulse:

\[
o(x, y, z) = \sum h(x-x', y-y', z-z') f(x', y', z')
\]
3.3. Determining the PSF of the Dynamic Data Sequence

\[ o(x, y, z) = \int \int \int f(x', y', z')h(x - x', y - y', z - z')dx'dy'dz' + n(x, y, z) \] (3.1)

where \( n(x, y, z) \) represents additive noise.

However, various factors conspire to complicate this idealised analysis. In particular, real MR scanners exhibit non-linearities from magnetic field inhomogeneities (Lewis and Fox, 2004), signal intensity to concentration relationships (Pedersen et al., 2004) or the effects of the magnitude operators used to produce images in the positive real domain (Steckner et al., 1994). However, some non-linearities can be overcome by suitable transformations or minimised by careful selection of parameters. Thus, although the linear shift-invariant imaging system is an idealised concept and is, in practice, never realised, some MR sequences can be considered sufficiently shift-invariant to be commonly characterised by their PSF (Steckner et al., 1994; Miyati et al., 2002).

However, it is usually very difficult to directly measure the PSF, as by definition, a point source must be small and the acquired PSF will only contain a few pixels with low contrast and might be affected by random noise. In practice, an estimate of the PSF, the Line Spread Function (LSF) is commonly used.

Figure 3.2: Plot of \( 1/T_1 \) relaxation against [Gd]
3.3.2 Line Spread Function

The LSF is a 1D approximation to the PSF, as it is an integrated profile of the PSF:

\[
\text{LSF}(x) = \int_{-\infty}^{\infty} \text{PSF}(x, y, z) \ dy \ dz
\]  

(3.2)

Thus, the LSF can be directly calculated from the PSF. However, the PSF cannot always be calculated from the LSF as the PSF contains information about the spatial resolution in all directions whilst the LSF is limited to only one specific direction. A system has only one PSF, but an infinite number of LSFs, one for each angle. Thus, measuring the LSF at a single angle does not provide enough information to calculate the complete PSF except in the case where the PSF is circularly symmetric. In cases where the PSF is not circularly symmetric, estimates for the PSF may be obtained from multiple LSFs measured at various angles. The PSF can then be constructed in a similar manner as images are in computed tomography through projection methods.

3.3.3 Line Spread Function Estimation

The LSF is itself usually estimated (rather than measured directly, as the same problems encountered on imaging a point are applicable to imaging a line) from an edge. A high contrast edge is commonly captured from phantom data to obtain an edge profile, or Edge Spread Function (ESF), the derivative of which is the LSF. Thus, for an edge profile \( \text{ESF}(x) \):

\[
\text{LSF}(x) = \frac{d}{dx} \text{ESF}(x)
\]  

(3.3)

In a discrete system, differentiation can be performed using finite-element differences so that:

\[
\text{LSF}(x_j) = \frac{\text{ESF}(x_j) - \text{ESF}(x_{j-1})}{x_j - x_{j-1}}
\]  

(3.4)

where \( j \) represents the sample number.
Phantom data were acquired using the VIBE sequence described Section 3.1. The phantom used had solid square perspex blocks whose boundaries provide high contrast edges for ESF estimation. These edges are located far from other structures and thus avoid ringing effects from adjacent structures (McRobbie, 1997). MR-captured phantom data can be seen in Figure 3.3.

**Figure 3.3:** MR image from phantom used for LSF estimation. The boundary of the square perspex blocks provides a high contrast edge for the extraction of ESFs. (Left) was acquired using the VIBE sequence but at a higher in-plane resolution (0.5x0.5mm) than that of the proposed clinical resolution (right).

In practice, the main issues to consider in estimating an LSF from an ESF are those of noise arising from a variety of sources. The differentiation step in Equation 3.4 is highly sensitive to noise and intensity variations in the ESF affect the LSF (intensity noise). Additionally, there is typically noise associated with large voxel sizes relative to the width of an edge, and therefore poor sampling of the edge profile (sampling noise). Another problem, specific to MR, is the effect of the magnitude operator applied after the IFT of the k-space data. These issues are addressed below.

**Effects of Noise**

The effects of noise on LSF estimation may be reduced by combining several adjacent ESFs. On a perfect sampling grid, with an ESF that is at exactly 90° to the angle, samples could be simply averaged to reduce intensity noise. However, sampling noise would still apply and would be strictly limited by voxel size. In practice, there is normally an angle $\alpha$ between the ESF and the edge. This is rather an advantage that...
allows the reduction of sampling noise through a technique known as oversampling: by imaging an angled edge, the edge is effectively over-sampled by adjacent profiles (Judy, 1976). This is illustrated in Figure 3.4, where for a row profile \( R \), each adjacent row over-samples the profile at a distance \( N \times \Delta x \), where \( N \) is the row number relative to \( R \), and \( \Delta x \) is given by:

\[
\Delta x = h \times \sin(\alpha)
\]  

(3.5)

Thus, by knowing \( \alpha \) and the location of the edge, several ESFs can be registered together to produce an oversampled ESF with reduced sampling noise. There are many approaches in the literature to accurately calculate both the position and the angle of the edge, ranging from linear interpolation along the profile to Hough Transforms. For a survey of some of these methods see Samei et al. (2005).

Figure 3.4: Oversampling an edge. Example for a row profile: moving from row \( R \) to row \( R+1 \) is equivalent to moving along the profile by an amount proportional to the height and the sine of the angle with the vertical.

However, in the author's experience global estimates of \( \alpha \) do not result in accurate alignment of the ESFs due to local variations of the angle \( \Delta \alpha \) along the edge produced by the sampling grid. A different approach is to find the relative displacement, \( \Delta x \), between adjacent profiles ESF\(_j\) and ESF\(_{j+1}\). A simple approach is to find a \( \Delta x \) that minimises the area between two (interpolated) profiles:

\[
\arg\min_x \sum_i |\text{ESF}_j(x_i) - \text{ESF}_{j+1}(x_i)|
\]  

(3.6)
Thus, multiple ESFs and LSFs might be registered, as can be seen in Figure 3.5. The registered LSFs might then be interpolated into the same grid and averaged together to produce a smoother version, or more appropriately, registered ESFs may be interpolated into the same grid and averaged to produce a smoothed profile that is then differentiated. However, due to the relatively large voxel size of the dynamic sequence as defined on Section 3.1, the characterisation of the LSF was still considered too poor. The methodology for further reducing the effects of sampling noise on the estimation, and therefore modelling, of the LSF is described below.

![Figure 3.5: LSFs from oversampled profiles registered and interpolated into the same sampling grid. The corresponding averaged LSF can be seen in blue.](image)

**Effect of Poor Sampling**

The temporal constraints to adequately sample renal function result in the need for very rapid acquisition of data and consequently a relatively low resolution. Even though edge profiles may be oversampled as described in the previous section, the resulting ESFs might still contain few samples and offer a poor characterisation of the profile and therefore of the LSF. This is evident from Figure 3.6 (right), where the lack of data prevents an accurate view of the underlying LSF. In the next section, a model for the LSF is proposed from the experimental data. The model was derived from data acquired using the same acquisition sequence but with increased spatial resolution. The equivalent LSF, acquired with an in-plane voxel dimensions of 0.5x0.5 mm, may
be seen in Figure 3.6 (left). The validity of the model for the lower resolution (voxel size 3.1x3.1 mm) is tested through the acquisition of data with several in-between resolutions and fitting of the model to the corresponding LSFs. These results may be seen in Section 3.3.4.

![Figure 3.6: Effects of poor sampling on LSF estimates: with a higher spatial resolution (left), but otherwise using the same sequence, the LSF is captured in more detail than using the proposed clinical acquisition sequence (right).](image)

**Effect of the Magnitude Operator on LSF and ESF estimates**

In theory, MR images should theoretically be real (Steckner et al., 1994), but negative and complex values in the IF-transformed data arise due to noise in k-space data. Thus, a magnitude operator is applied to the IF-transformed data that, for the purpose of visualisation, transforms these image data voxel values into natural numbers, \( \mathbb{N}_0 \).

The phantom used to estimate the LSFs contained perspex blocks that produce no MR signal, within an aqueous solution that produces a large MR signal, with the transition between these two compartments defining the edge. The effects of the magnitude operator is clearly visible in the zero-signal portion of the edge profile (see Figure 3.7). Here, any boundary effects associated with the LSF are essentially cropped as negative values are converted into the positive domain. This prevents observation of the true characteristics of the underlying LSF (for non-zero signal). A solution to this could be to use a phantom with boundaries defined by a low-signal/high-signal edge, instead of...
3.3. Determining the PSF of the Dynamic Data Sequence

a zero-signal/high-signal edge. However, in our case, the fitting range of the model was limited to exclude the zero-signal part of the profiles. Thus:

\[
\text{LSF}'(x_i) = \begin{cases} 
\text{LSF}(x_i) & \text{for } x_i < x_m \\
0 & \text{otherwise}
\end{cases}
\]  

where \(x_m\) is defined as \(\arg(\max \text{LSF} - \min \text{LSF})/2\)

\[\text{(3.7)}\]

Figure 3.7: Effects of Magnitude operator on LSF. The magnitude operator transforms the input data into \(N_0\). This affects low-intensity voxels and as it can be clearly seen in the parts of the LSF that correspond to low-intensity in the ESF, this creates an asymmetric LSF.

3.3.4 Empirical LSF Model for the DCE-MRI sequence

The specific shape of the PSF of a particular MR sequence is scanner-dependent as it is influenced by several factors. Possibly, the two most important may be those associated with finite sampling and MR signal decay.

The sampling of \(k\)-space during image acquisition results in a finite set of spatial frequencies being captured. Thus, it is common in the literature to consider the PSF of a MR scanner as a Sinc function (Miyati et al., 2002; Links et al., 1998; Fain et al., 1999; Sato et al., 2003), that is, the inverse Fourier Transform of an ideal low-pass filter. However, in addition to this, there are further filtering effects due relaxation within the temporal sampling window that result in exponential signal decay during acquisition.
In particular, GE sequences such as the VIBE sequence used here (Rofsky et al., 1999) are affected by $T_2$ and $T_2^*$ decay (Windischberger and Moser, September 2000; Haacke et al., 1999). To model these effects, a Sinc-Gaussian mixture, $\text{LSF}_{gs}(x)$, is proposed. Thus, in 1D:

$$\text{LSF}_{gs}(x) = h_s \times \text{sinc}\left(\frac{x}{\sigma_s} - \mu_{gs}\right) + h_g \times \exp\left(-\frac{(x - \mu_{gs})^2}{2\sigma_g^2}\right)$$

(3.8)

where $\mu_{gs}$, $\sigma_s$, $h_s$, $h_g$, and $\sigma_g$ are the fitting parameters.

Model Fit

LSF obtained from in-plane edges at $0^\circ$, $45^\circ$, $90^\circ$, and $135^\circ$. LSFs were also obtained from an edge perpendicular to the acquisition plane for the LSF in the $z$-direction. Figure 3.8 shows exemplar fittings for in-plane LSFs at $0^\circ$. As can be seen, the samples at the clinical resolution are insufficient to provide an accurate characterisation of the LSF. Figure 3.9 shows exemplar fittings for in-plane LSFs at $90^\circ$. 
3.3. Determining the PSF of the Dynamic Data Sequence

Figure 3.8: Examples of model fits to LSFs on 0° edges for three voxel sizes: (top) 0.5x0.5 mm, (middle) 1.5x1.5 mm, and (bottom) 3.2x3.2 mm.
Figure 3.9: Examples of model fits to LSFs on 90° edges for three voxel sizes: (top) 0.5x0.5 mm, (middle) 1.5x1.5 mm, and (bottom) 3.2x3.2 mm.
The variability of the FWHM of the fitted LSFs as a function of spatial location and orientation is investigated by analysing the variability in FWHM estimates for edges with different orientations (in plane at 0°, 45°, 90°, 135°, and out of plane, z-direction) and at different locations and through different image slices of the acquired phantom.

**3D Point Spread Function**

The average FWHM as a function of voxel dimension for the in-plane PSF can be seen in Figure 3.10. As can be seen, the uncertainty on the FWHM estimates increases as the voxel dimension increases.

![Figure 3.10: FWHM variability as a function of voxel dimension. There is a linear relationship between the FWHM and voxel dimension](image.png)

In order to produce accurate estimates of the model parameters at the voxel dimension we are interested in (3.2x3.2 mm), the smaller voxel dimension estimates might be used: The relationship between the model parameters $\sigma_s$, and $\sigma_g$ and voxel dimension can be seen in Figure 3.11 and 3.12 respectively. This linear relationship suggests that the Sinc-Gaussian model may be accurately fitted to the smaller voxel dimension data and the parameters at the clinical voxel dimension obtained by interpolation of the higher smaller voxel dimension parameters.

Figures 3.11 and 3.12 show the model fitting becomes inaccurate at the voxel dimensions of the clinical sequence (3.2x3.2 mm). In particular, the Sinc component dominates and the Gaussian component is difficult to model ($\sigma_g$ is off the chart in Figure 3.12).
Figure 3.11: Variability of model parameter $\sigma_s$ as a function of voxel dimension. As with the FWHM, there is a linear relationship between the parameter estimates and voxel dimension. The last point (3.2 mm) is an outlier resulting from poor fittings of the model being dominated by the Sinc component due to noisy and insufficient number of samples available at this voxel size (see Figure 3.8-bottom).

Figure 3.12: Variability of model parameter $\sigma_g$ as a function of voxel dimension. As with the FWHM and $\sigma_s$, there is a linear relationship between the parameter estimates and voxel dimension. In this case, the last point (3.2 mm) is also an outlier (falls outside the range shown in the graph) resulting from poor fittings of the model. This justifies the extrapolation of the higher resolution parameters.

Therefore, the PSF parameters for the model in the next Section were estimated from the higher voxel dimension scans and extrapolated to the (3.2x3.2 mm) case.
3.4. Movement Correction

Based on the model fits to the LSF data in the $x$, $y$, and $z$ directions a 3D anisotropic PSF was produced:

$$\text{PSF}(x) = A \times \left( h_s \times \text{sinc} \left( \frac{x}{\sigma_s} \right) + h_g \times \exp \left( -\frac{x^2}{2\sigma_g^2} \right) \right)$$  \hspace{1cm} (3.9)

where $h_s = [h_x, h_y, h_z]^T$, $h_g = [h_x, h_y, h_z]^T$, $\sigma_s = [\sigma_x, \sigma_y, \sigma_z]^T$ and $\sigma_g = [\sigma_x, \sigma_y, \sigma_z]^T$ determine the height and width of the Sinc and Gaussian components respectively, $x = [x, y, z]^T$ and $A$ is a normalising factor.

Thus, after fitting the model to the high resolution phantom data and extrapolating to the clinical voxel-dimensions, the resulting PSF for the DCE-MRI sequence was a Sinc-Gaussian model with in-plane FWHM = 6.3 mm ($x$) and 6.4 mm ($y$) and out-of-plane FWHM = 14.8 mm.

3.4 Movement Correction

In this section, the implementation of a 3D movement correction algorithm for DCE-MRI data is presented. The approach follows the 2D of Giele et al. (2001), discussed in Section 2.3.4. This approach presents two characteristics that make it a suitable for developing into a fast and robust 3D movement correction method. First, it requires only a single segmentation of the kidney to create a mask which effectively limits the search space. Secondly, the approach allows the estimation of translation in a fast and non-iterative manner in the Fourier domain.

3.4.1 Method

The proposed approach is based on the creation of a grayscale mask that is used to localise the search to the area surrounding the kidney. After applying the mask to consecutive image slices, an estimate of the shift between kidneys can be obtained by the difference in the phase images of the FT from the masked images. The shift is then used to displace the search volume accordingly, which then becomes the reference for the next volume. This movement correction method is further described below.
Chapter 3. MR Data Acquisition and Preprocessing

Mask Generation

The use of a mask in this approach is motivated by the presence of other objects in the MR image, such as the liver or the other kidney, which may present different movement profiles and therefore obfuscate movement estimation for the kidney. The purpose of the mask is therefore, to eliminate the effects of any background clutter, whilst still covering all possible movements that the kidney might have experienced during the scan.

Initially, a binary mask, \( M_b(x) \), is created from a 3D kidney contour, \( C(x) \) where every voxel within the contour is set to 1. In our case, the kidney contour was manually segmented from a high-contrast volume. \( M_b(x) \) is then extended in all directions to create an enlarged binary mask \( M'_b(x) \), assumed to cover all of that kidney's movement due to breathing. Therefore the main enlargement is in the head-feet direction. This is accomplished through a filtering step, using an anisotropic Gaussian filter. A multivariate \( d \)-dimensional gaussian distribution is defined as:

\[
H(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp \left( -\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) \right)
\]  

(3.10)

where \( x \) is a vector of \( d \) random variables, of which \( \mu \) is the mean vector and \( \Sigma \) is the covariance matrix. Equation 3.10 may be used to create a normalised filter:

\[
H_{\text{norm}}(x) = \frac{H(x)}{\sum_x H(x)}
\]  

(3.11)

As per Giele et al. (2001), standard deviation values of 5 and 20 pixels in the horizontal and vertical (head to feet) directions respectively were used for the filter. For this 3D extension, a standard deviation of 5 was used in the voxel-thickness dimension. The result of this filtering using the anisotropic filter is a grayscale mask \( M(x) \):

\[
M(x) = H_{\text{norm}}(x) * M_b(x)
\]  

(3.12)

which can then be thresholded to obtain \( M'_b(x) \):

\[
M'_b(x) = \begin{cases} 
M(x) & \text{for } M(x) > T \\
0 & \text{otherwise} 
\end{cases}
\]  

(3.13)
3.4. Movement Correction

where $T$ controls the degree of enlargement. However, $M_b'(x)$ cannot be used directly as a mask since movement correction techniques would attempt to align the strong edges of this mask. Thus, a second Gaussian filter, $H_2$, is applied, using standard deviation values of 10 pixels in all dimensions, for the purpose of softening mask edges, resulting in the grayscale enlargement mask $M_g(x)$:

$$M_g(x) = H_2(x) * M_b'(x) \quad (3.14)$$

Finally, the mask is applied to both the reference and search volumes to give the masked volumes $V_f'$ and $V_s'$:

$$V_f'(x) = V_r(x) \times M_g(x)$$
$$V_s'(x) = V_s(x) \times M_g(x) \quad (3.15)$$

Shift Estimation

Shift estimation is performed in frequency space, using the search $\Phi(Y_s)$ and reference $\Phi(Y_r)$ 'phase' volumes:

$$\Phi(Y_s) = \arctan \left( \frac{\text{Im}(Y_s)}{\text{Re}(Y_s)} \right)$$

$$\Phi(Y_r) = \arctan \left( \frac{\text{Im}(Y_r)}{\text{Re}(Y_r)} \right) \quad (3.16)$$

where $Y_s$ and $Y_r$ are the Fourier-transforms of the search and reference masked volumes respectively:

$$Y_s(w) = \mathcal{F}[V_f'(x)]Y_r(w) = \mathcal{F}[V_r'(x)] \quad (3.17)$$

The phase difference $\Psi$ is then found as:

$$\Psi(Y_s, Y_r) = \Phi(Y_s) - \Phi(Y_r) \quad (3.18)$$

The phase difference volume is then combined with a flat magnitude $M(w) = \text{constant}$, and the inverse Fourier-transform is applied:

$$Z(x) = \mathcal{F}^{-1} \left[ M(w) \times e^{i\Psi(Y_s, Y_r)} \right] \quad (3.19)$$

The resulting 'image space' volume contains a peak at location $x_0$, corresponding to the 3D displacement between the two volumes.
3.4.2 Results and Discussion

The 3D movement correction technique was applied to two sets of simulated data generated using the Synthetic Data Simulator that is described in the next chapter. Synthetic kidney volumes were created using mean intensities and variances obtained from clinical MR Renography data. Exemplar image slices of the generated synthetic data can be seen in Figure 4.4. Three different datasets were created. The first (a) contained no significant (only noise) intensity changes between consecutive data volumes, typical of pre-contrast data. The second dataset (b) contained large intensity changes, typical of those associated with the blood perfusion phase. Finally, a third dataset (c) contained smaller and gradual intensity changes, such as those associated with the filtration and secretion and re-absorption phases of kidney function.

Table 3.1 shows results for movement correction on these synthetic data. It also shows movement correction results on clinical data (see Section 3.1 for details on the acquisition), where the GT was obtained by an operator manually placing the segmented kidney contour on the (visually) correct location for a series of key data volumes on two volunteer datasets at each of the three intensity change stages specified above.

Table 3.1: Movement correction results for MR Renography data corresponding to the pre-contrast phase (a), blood perfusion phase (b) and filtration and secretion and re-absorption phases (c).

<table>
<thead>
<tr>
<th></th>
<th>synthetic</th>
<th>clinical</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>exact ±1</td>
<td>exact ±1</td>
</tr>
<tr>
<td>Dataset a</td>
<td>100 100</td>
<td>76 91</td>
</tr>
<tr>
<td>Dataset b</td>
<td>96 98</td>
<td>71 84</td>
</tr>
<tr>
<td>Dataset c</td>
<td>100 100</td>
<td>62 77</td>
</tr>
</tbody>
</table>

The generated synthetic movement ranged from 0 to 10 voxels in the head-feet direction, 0 to 5 in the left-right direction, and 0 to 3 voxels in the anterior-posterior direction. As can be seen, whole-voxel translations were accurately estimated on the synthetic data, demonstrating the validity of the method. Errors were only found on displacements
larger than 8 voxels, due to the use of a static mask that is therefore positioned on the edges of kidneys that are significantly displaced.

For the clinical data, the largest movement present (from the GT) was found to be 4 voxels. The results are worst for the volumes with intensity changes (Datasets b and c). In particular, the performance is worst for the data volumes in the secretion and re-absorption phase. This is probably explained by the larger changes in tracer distribution as it moves in and out of the medulla and into the renal pelvis. In contrast, the performance on Dataset (b) is slightly better. Even though the intensity changes are largest in this interval, the tracer is mostly distributed around the outer cortex, providing a strong edge of the kidney contour which helps to accurately locate the organ. Overall correct estimation results were 67%, with 80% of the cases correctly estimated to within 1 voxel movement. The results on this 3D extension are similar to the 2D work of Giele et al. (2001).

However, it is important to note that the proposed method does not estimate rotational movement or sub-voxel displacement. Methods that do take into account these effects may produce smoother renograms and therefore more accurate renal function estimates. However, it may be argued that increased accuracy in movement correction is not the most important factor in the analysis of MR Renography data. There are other factors, such as the influence of PVs that may prove more important factors in the development of robust MR Renography techniques. Therefore, the rest of this thesis is concerned with addressing these effects.

### 3.5 Conclusions

In this Chapter details of the data acquisition, both for the DCE-MRI and anatomical data used in this work have been presented. The $T_1$-weighted DCE-MRI sequence captures volumetric datasets covering the whole abdomen every 2.5 s. The anatomical scan, acquired pre-contrast, acquires the same FOV into a higher resolution grid. Dynamic data voxel dimensions were chosen as integer multiples of the anatomic data voxel dimensions to minimise registration errors.
An empirical model for the PSF of the DCE-MRI sequence is then presented. This model was obtained from phantom data by analysing LSFs in $x$, $y$ and $z$ directions. A Sinc-Gaussian model was proposed and fitted to the phantom data at various voxel dimensions, demonstrating a linear relationship between the model parameters and voxel size. This enabled accurate estimation of the model parameters at high resolution which then were extrapolated to the more noisy voxel-dimensions used for the clinical data. Therefore, a 3D model of the PSF for DCE-MRI sequence was obtained. This model is used in the next few chapters as the basis of the synthetic data simulator and the two PV correction techniques that are proposed.

The last section of this chapter is concerned with movement correction of the DCE-MRI data. A 3D movement correction method was proposed. Results on clinical data showed that translational movement was correctly estimated for 63% of the cases, with 80% of the movement correctly estimated to within 1 voxel when compared to manually generated GT.
Chapter 4

Synthetic Data Simulator

In Chapter 2, a review of the most important issues in Magnetic Resonance (MR) Renography, and the analysis and processing steps that have been applied to address them, was presented. In particular, it was highlighted how the lack of available Ground Truth (GT) is an impediment in the assessment of movement correction techniques. Similarly, it is difficult to obtain appropriate GT data for capturing the Partial Volume (PV) effect in physical studies. In fact, these problems are not only experienced in MR imaging, but across the biomedical imaging field.

In this chapter, the design and implementation of a 4D (three spatial dimensions plus time) synthetic data generator developed for simulating the effects of varying position, noise, and partial volumes in dynamic data, is presented. The purpose of the simulator is to create synthetic data, with its inherent GT, suitable for the assessment and development of image processing techniques to address these effects. Even though the simulator has been developed for MR data, the generic approach used (see Section 4.1) makes it possible to generate simulated data from other imaging modalities (e.g. Section 4.7.2 shows the simulation of a Positron Emission Tomography (PET) phantom).

Section 4.2 presents an overview of the simulator, and its main modules/components are described in Sections 4.3 to 4.6. Exemplar simulated data, is presented in Section 4.7.
Chapter 4. Synthetic Data Simulator

4.1 Simulated Data in Medical Imaging

A key issue in the development of image processing techniques for all medical imaging applications is how to evaluate and validate the proposed methodology. Objective validation of automated image processing techniques, such as movement correction or segmentation, applied to in vivo studies is difficult, lacking GT reference data. Many studies address this issue by comparison to manually generated results from expert operators. However, this approach may suffer from intra- and inter-expert variability and is, in many cases, too labour intensive for extensive assessment. Another approach, specially used in functional applications, is that of assessing the goodness of fit of the results to some idealised expected model as a surrogate measure of the algorithm's performance. A further problem of in vivo studies is the difficulties to reproduce results, which in many cases are scanner-dependent and affected by a wide range of variables.

Ground-truth reference data may be obtained by the use of physical phantoms of known geometric and magnetic resonance properties. However, it is very difficult and time-consuming to design realistic mechanical phantoms that model the anatomical and functional properties of even the simplest organ.

Within this context, computer-generated simulated data presents an interesting approach in terms of designing more complex anatomical structures exhibiting more realistic functional characteristics than physical phantoms. Furthermore, simulated data is reproducible and provides ground truth information on which to base an objective assessment of applied image processing techniques.

Simulated data is based on a set of assumptions and idealised models that often result in a best-case scenario representing a simplified model of the variability associated with true data. Thus, the user must be cautious about extrapolating results from tests on simulated data, and any technique must ultimately be validated on in vivo studies. However, simulated data may be highly valuable in the evaluation stages, and may be used to objectively compare the relative performance and accuracy of competing algorithms to a sometimes necessary precision not always possible in real physical studies.
4.1.1 Generation of Synthetic MR Data

There are various approaches to generate synthetic MR image data through computer simulation in the literature, especially in the field of brain imaging. These approaches differ mainly in the closeness to reality of the resulting data and the computational cost necessary to achieve it.

One of the first MRI simulators was proposed by Bittoun et al. (1984) based on numerically solving the Bloch Equations at every point of discrete 1D object data. This approach results in highly realistic synthetic data capable of simulating many of the characteristics of real MR data, but requires a high computational cost that becomes impractical when applied to higher dimensional data (Olsson et al., 1995). Recent work has focused on implementations of this kind of approach through distributed grid architectures (Benoit-Cattin et al., 2005).

A different approach to generating synthetic data, at low computational cost, is to use tissue templates. These are spatial masks representing different tissues to which tissue-dependent intensities and noise are assigned (Collins et al., 1995; Herndon et al., 1998). However, images synthesised in this manner are more simplistic and unable to simulate many of the artefacts associated with the MR image formation process. There are also hybrid approaches that use tissue templates and associate to them a signal intensity derived from discrete-event Bloch Equations (Kwan et al., 1999).

From the above discussion, it should become clear that the simulation of synthetic MR data must be driven from the particular effects to be investigated. In this work, the interests are in simulating 4D dynamic data within a relatively large FOV, affected by the PV effect, organ movement and variable signal intensities and noise as a result of the passage of tracer.

Large datasets consisting of series of 3D volumetric simulations are required for generating movement artefacts. Generating PVs may be accomplished by downsampling high resolution data into a lower resolution grid. The combination of these two effects results in high computational cost to produce multiple large-size 3D volumetric simulations. Therefore, the simpler template-based approach was considered appropriate for this work.
On a template-based approach, high-resolution templates represent an approximation to real world physical objects, unconstrained by digitisation or bandwidth. These constrains are introduced by the imaging system, and are generally defined by its PSF. Therefore, this approach has the additional advantage of being able to model data, not only from MRI but, from other acquisition systems such as a PET scanner, by the use of an appropriate PSF.

4.2 A Synthetic Data Generator

The proposed template-based synthetic data simulator consists of four main steps. These are the creation of 3D organ and tissue templates, addition of movement, addition of tissue and/or tracer models, and downsampling. A diagram of the design can be seen in Figure 4.1.

Initially, high resolution 3D volume templates are produced consisting of labelled data corresponding to the required physical structures. For example, in the case of DCE-MRI data for Renography, typical abdominal structures would correspond to the aorta, liver, kidneys, etc., and to the tissues within these organs/structures (such as renal cortex, medulla and pelvis for the kidneys).

A movement model, if required, may be introduced at this stage. Displacement of the high resolution data can be used as GT for the assessment of sub-pixel performance of movement correction techniques applied to the synthetic low resolution data. Further, a movement model may also be used to test the influence of small movement in PVE correction techniques.

The next stage is the allocation of tissue signal intensity values to the different compartments according pre-defined models. This allows modelling the passage of tracer through each compartment model. Simulated noise is also introduced at this stage.

The final step in the process is the downsampling of the high resolution templates into the desired lower resolution, thus creating the PV effect on the down-sampled synthetic low resolution data. Downsampling consists of filtering and decimating the
4.3. High Resolution Templates

Figure 4.1: Flow diagram for a synthetic data generator. The resulting low resolution data is intended to simulate the PV effect and movement artefacts of contrast enhanced dynamic data.

The resulting low resolution data in a manner that closely relates to the low-pass filtering and sampling effects produced by the PSF of the acquisition system.

The resulting low resolution data represents the scanner data, band-limited by the PSF and subject to noise and other artefacts.

4.3 High Resolution Templates

This section describes the generation of discrete high resolution 3D volume templates of labelled data for the kidneys and other abdominal structures that aim to represent
the continuum case. The approach is based on defining a number, \( n_O \), of objects \( O_i \), for \( i = 1, ..., n_O \), as rigid volumetric templates for each organ. Each object is in turn composed of a number, \( n_C \), of compartments \( C_{ik} \), \( k = 1, ..., n_C \), corresponding to the different tissues within the organ. Thus, each voxel \( \omega \) is assigned a label \( C_{ik} \) representing the \( k \)-th compartment (tissue) of the \( i \)-th object (organ).

Two possible approaches to generating such templates are the use of 3D geometrical shapes and the segmentation of high resolution anatomical data. Whilst non-anatomically correct, it is arguable that simple geometrical shapes might be sufficient to estimate macroscopic effects such as organ movement. For movement correction techniques that are based on intensity similarity measures (i.e. not estimating movement based on a specific shape model) the shape of the structures should not matter.

### 4.3.1 3D Geometric Shapes

Within the simulator, 3D geometric shapes are defined using superellipsoids (Barr, 1981). Superellipsoids are a family of closed quadric surfaces that can represent many distinct shapes. Superellipsoids can be defined by the following implicit equation:

\[
\left( \frac{x}{a_x} \right)^\frac{2}{\epsilon_2} + \left( \frac{y}{a_y} \right)^\frac{2}{\epsilon_1} + \left( \frac{z}{a_z} \right)^\frac{2}{\epsilon_2} = 1 \tag{4.1}
\]

where \( a_x, a_y, \) and \( a_z \in \mathbb{R}^+ \) are the axial lengths, and \( \epsilon_2 \) and \( \epsilon_1 \) control the shape of the surface parallel to the \( x-y \) plane and on a plane perpendicular to the \( x-y \) axis that contains the \( z \) axis respectively.

For the construction of volumetric templates, an inside-outside function \( F(x, y, z) \) may be defined such that:

\[
F(x, y, z) = \left( \frac{x}{a_x} \right)^\frac{2}{\epsilon_2} + \left( \frac{y}{a_y} \right)^\frac{2}{\epsilon_2} + \left( \frac{z}{a_z} \right)^\frac{2}{\epsilon_2} \tag{4.2}
\]
where, for \( F > 1 \), the point \((x, y, z)\) lies outside the superellipsoid, and for \( F \leq 1 \), the point \((x, y, z)\) lies on the surface or within the surface and is thus part of the volume. By varying the value of \( \epsilon_1, \epsilon_2 \) and the axial lengths, a large variety of 3D shapes may be generated (see Figure 4.2).

![Figure 4.2](image)

**Figure 4.2:** 3D renderings of superellipsoid surfaces obtained by varying \( \epsilon_1 \) and \( \epsilon_2 \) in the range \( 0.1 < \epsilon_1, \epsilon_2 < 2 \).

Equation 4.2 may be used to define a compartment within a local coordinate system. For every compartment \( C_{ik} \) belonging to object \( O_i \), its 3D pose (location and orientation) within the object space is defined by \( \Lambda_{ik} \) so that any point in compartment space, \( x = [x, y, z]^T \in C_{ik} \), is transformed into object space as \( x' \) by:

\[
x' = \Lambda_{ik} x = R_{ik} x + t_{ik}
\] (4.3)
where $R_{ik}$ is a rotation matrix\footnote{Rotations are specified through Euler angles using the following convention, $O_i$ is first rotated by an angle $\phi$ about the x-axis, followed by a rotation $\theta$ about the y-axis, and finally a rotation $\psi$ about the z-axis. Within this convention, all rotations are defined as anticlockwise and about the original axis.} with respect to $S_{ik}$’s centre of mass and $t_{ik}$ is a translation with respect to the origin of each object’s space, that is $O_i$’s centre of mass.

For convenience, $\Lambda_{ik}$, can be expressed in homogeneous coordinates as:

$$\Lambda_{ik} = \begin{bmatrix} R_{ik} & t_{ik} \\ 0^T & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & tx_{ik} \\ r_{21} & r_{22} & r_{23} & ty_{ik} \\ r_{31} & r_{32} & r_{33} & tz_{ik} \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.4)$$

hence, rotation and translation are combined into a single operation and successive transformations can be concatenated.

Once an object $O_i$ has been defined in terms of its constituent compartments $C_{ik}$, each object’s 3D pose is defined within an overall FOV space by $L_i$ so that any point $x' \in O_i$ is transformed into $x_a$ by:

$$x_a = L_i x' = L_i \Lambda_{ik} x \quad (4.5)$$

where $L_i(tx_i, ty_i, tz_i, \phi_i, \theta_i, \psi_i)$ is defined in the same manner as $\Lambda_{ik}$ above, but with the parameters specified within the ‘object definition’ for $O_i$. Thus, the FOV space may be defined containing a series of objects, each composed of a subset of compartments.

### 4.3.2 Anatomical Data Segmentation

Another method for generating high resolution templates is through the segmentation of anatomical structures from high resolution data acquired with, for example, MRI or CT. 3D organ boundaries might be obtained by segmentation and converted to tissue templates that can be combined to form an abdominal map. The resulting volumetric 3D templates may be used as an input to the synthetic data generator. However, this
approach relies on a method for automated 3D or 2D segmentation of high resolution anatomical data as manual segmentation is likely to be too labour-intensive. Further, a method for generating isotropic 3D data from a set of or 2D slices might also be required.

4.4 Movement Model

To simulate the effects of organ movement, object-specific movement models may be added to the high resolution templates. For every object $O_i$, a movement model $M_i(t)$ is defined, where $t$ is the time (in seconds) elapsed from the beginning of the sequence. The movement model specifies a linear transformation of six degrees of freedom that is added to each object's intrinsic transformation $L_i$ (note that in the case of organ templates generated from anatomical data segmentation, $L_i$ is the identity matrix $I$ since it is assumed that the segmented templates are spatially already located correctly). Thus, for each object $O_i$ at time $t$, a new linear transformation $L'_i(t)$ is used instead of $L_i$ to denote the object's position and orientation:

$$L'_i(t) = L_i + M_i(t) \quad (4.6)$$

The inclusion of a movement model allows assessment of the performance of movement correction schemes on simulated data. Movement correction on down-sampled data allows assessment of the response of these methods to sub-pixel translations and rotations induced on the high resolution data. Furthermore, these movement models may also be used to investigate the effects of sub-voxel movement on the PV effect.

4.5 Tissue and Tracer Models

This section describes the assignment to every voxel $\omega$ within each 3D high resolution volume of an intensity value. Within the template definition described in Section 4.3, a tissue type is set for every compartment $C_{ik}$. The synthetic data generator then assigns
intensity values to every voxel according to its tissue type, a tracer dynamics model, and required noise.

4.5.1 Tissue and Tracer Intensity Models

Within the generator, every tissue type has a tissue and tracer intensity model $T_j(t)$ associated to it, where $j = 1, \ldots, n_T$ ($n_T$ is the number of tissues), so that:

$$T_j(t) = I_j \times (1 + E_j(t))$$

(4.7)

where $I_j$ represents a tissue-specific intensity model and $E_j(t)$ is the enhancement that may or not be produced in each tissue by tracer passage.

The simplest method for building such a tissue and tracer model applicable to particular scanner, acquisition sequence, tracer type and dosage, is through ROI analysis of volunteer data. Mean tissue intensity values may be obtained from ROIs placed on the relevant tissues in dynamic data, acquired with the required protocol. However, it should be stressed that careful ROI definition is required: placing ROIs within $\approx 1 \times$ FWHM distance of any boundary is likely to corrupt the resulting time-intensity plot due to spill-over from adjacent tissues (i.e. the PV effect). Thus, $I_j$ and $E_j(t)$ may be estimated respectively from samples acquired before and after the arrival of tracer to each tissue. This would result in a discrete $T_j(t)$, its sampling dictated by the acquisition time of the acquisition series. However, it would be possible to obtain from these data continuous functional models of tracer kinetics, such as in the renal data example in Section 4.7.1.

4.5.2 Noise Models

Associated with the mean intensity for each tissue in the model, there is also a variance that reflects inhomogeneities that might be present in the acquired data. This noise may be due to a variety of sources. Some variability will arise from intrinsic physiological tissue variation, whilst additional noise may also be induced by different components of the acquisition system, such as the electronics and processing that will
produce sampling and quantisation. Within the generator, these factors are grouped together as a single noise model.

For the case of MR data, the noise model used is a Gaussian distribution. MR magnitude image data is the result of the Inverse Fourier Transform on data acquired in $k$-space. This results in complex data in image space to which, for visualisation, a magnitude operator is usually applied. Assuming that noise in the real and imaginary data is Gaussian distributed, this results in Rayleigh-distributed noise in the magnitude image data (Henkelman, 1985; Gudbjartsson and Patz, 1995). However, the Rayleigh distribution becomes symmetrical and may be approximated by a Gaussian distribution for large SNR (for SNR values 5 skewness falls below 0.1 (Chung and Noble, 1999)).

In other modalities (those using photon counting: X-ray and radioisotope imaging), Poisson noise may be required at low SNR. The Poisson model has not been implemented here, however, as in the MR case, the noise distribution in these modalities also approximates a Gaussian as signal rises.

4.5.3 Generation of Synthetic Gaussian Noise

Synthetic Gaussian noise may be generated by using a Pseudo-Random Number Generator (PRNG), where a sequence of numbers is deterministically obtained, whilst giving the appearance of being a random sequence. A common approach to generating a sequence of random numbers with a particular distribution is to generate first a Uniform distribution $U$, and then transform it to the desired distribution. A more detailed description of random number generation and the generation of Gaussian distributed noise may be seen in Appendix D.

Uniform Distribution

The Uniform distribution for the proposed synthetic generator is obtained by the use of the Mersenne Twister PRNG, proposed by Matsumoto and Nishimura (1998). The Mersenne twister algorithm is a widely used random number generator for statistical simulations, due to the large period it can achieve and the small inter-correlation of the
generated data samples that it has achieved when extensively tested. In generating a
time series of synthetic data, the PRNG is initialised for every generated volume, which
thus avoids using the same random number for every voxel within the time sequence.
A possible approach to initialise a PRNG is to use the ‘time’ function of the computer.
However, this type of approach is unsuitable as the seed values are un-reproducible and
may be a correlated set of seeds due to processor granularity. The proposed approach
is to use a table of random number seeds according to the required time-stamp.

Gaussian Distribution

To generate a Normal distributed sequence \( N(\mu, \sigma^2) \) with zero-mean and unit-variance
of random numbers from a uniform sequence \( U \), the Box-Muller transform is used (Box
and Muller, 1958). Then, Gaussian distributed random variables \( X \) with arbitrary
mean and variance can be easily generated using the following property:

\[
\text{if } X \sim N(\mu, \sigma^2) \text{ and } a \text{ and } b \text{ are real numbers, then } aX + b \sim N(a\mu + b, (a\sigma)^2)
\]

(4.8)

Thus, within the generator, every tissue type has a noise model \( V_j(t) \) associated to it,
where \( j = 1, ..., n_T \) (\( n_T \) is the number of tissues).

4.6 Downsampling

The generation of low-resolution data volumes from high resolution data is known
as downsampling. Downsampling typically involves two steps. A low-pass filtering
step that removes high frequencies and a decimation step that resamples the high
resolution data into the lower resolution grid. Perhaps the most intuitive method for
downsampling is the spatial domain block-averaging filter. With this approach, both
steps are merged into a single one where the down-sampled data is the mean of an
equally-weighted cubic grid of high resolution data. However, the block averaging filter
does not often characterise the filtering effects produced by imaging acquisition systems,
such as MR or PET scanners. Low-pass filtering and decimation are further discussed
below.
4.6. Downsampling

4.6.1 Low Pass Filtering

Image acquisition systems have a finite bandwidth which results in low pass filtering of the acquired data. The main filtering step in MR acquisition is performed on the frequency space as data are sampled into a finite grid, and depends on scanner-specific factors and on the particular acquisition sequence. It is difficult to determine theoretically the combined effects of this step and the further filtering due to relaxation of the MR signal, reconstruction algorithm, and magnitude operator on the complex raw data. However, if the system is modelled as linear and shift-invariant, an estimate of these combined effects can be obtained by empirically measuring its PSF.

A detailed description on estimating the PSF for the proposed MR sequence can be found in Chapter 3. This approach to modelling the effects of the acquisition system has the further advantage of enabling the use of arbitrary PSFs, which combined with appropriate tracer models may be used to simulate data acquired with other imaging systems such as PET. Thus the simulator filters data according to user-specified PSFs which could be based on Gaussian, Sinc, or other filtering kernels, and with a measure of the blurring specified by its FWHM in x, y, and z directions.

4.6.2 Effect of Filtering on Variance

An effect of low pass filtering on noisy data is the reduction of variance. As the noise is added to the high resolution unfiltered data, the filtering effects of the downsampling stage have to be considered. On a simple block averaging filter, the relationship between the filtered standard deviation $\sigma_{\text{ave}}$ and the unfiltered standard deviation $\sigma$ for a set of independent and time-invariant samples $n$ is given by:

$$\sigma_{\text{ave}} = \sigma \sqrt{\frac{1}{n}}$$

(4.9)

Thus, to obtain filtered data with a standard deviation $\sigma_{\text{ave}}$ the values of the unfiltered data have to be multiplied by $\sqrt{n}$. However, such a relationship may not be easily obtainable for other filters such as Gaussian or Sinc functions. Further, if the PSF of
the simulator is not determined analytical but experimentally, the relationship between
the unfiltered and filtered data has to be calculated empirically.

For a given PSF, the relationship between filtered $\sigma^{LR}$ and unfiltered $\sigma^{HR}$ variances
can be empirically plotted (see Figure 4.3).

![Figure 4.3: Filtering effects on variance: Plot of the reduction on variance from $\sigma^{HR}$ to $\sigma^{LR}$, filtered with a 3D isotropic Gaussian spatial filter with $\sigma^{PSF} = 2.2$ mm.](image)

The gradient, $m$, of this straight line is thus the required factor needed to obtain the
correct high resolution variance. Therefore:

$$\sigma^{HR} = \frac{\sigma^{LR}}{m} \quad (4.10)$$

An effect of this required transformation to the variance, is that due to the large
variances needed in the high resolution data, the range of intensity values necessary to
define tissue distributions becomes large and, in practice, usually containing negative
values. Intensity values below zero may be truncated and this would result in non-
Gaussian distributed noise. This is solved by biasing the whole dataset with an intensity
$B$, which is a multiple of the minimum intensity in any of the tracer models:

$$B \geq \min_{j,t} (T_j(t) - 3 \times \sigma_j^{HR}(t)) \quad (4.11)$$

where the factor 3 ensures that 99.9% of the samples in the Gaussian distributions are
covered. Once filtered, data have to be de-shifted by the same amount $B$, so that the
final distributions are centred on the required mean intensities, $T_j(t)$.
4.6.3 Decimating

Having performed the filtering step with a filter that approximates the PSF of the acquisition system, the next step is that of decimation. This simply involves re-sampling the data into the desired low resolution grid. User-defined decimating factors $d_f$ are used for each $x$, $y$, and $z$ direction. To avoid aliasing problems with the sampling, $d_f$ must be chosen so that:

$$d_f \leq \frac{\min_{i_k} \{\text{width}_{i_k}\}}{2}$$  \hspace{1cm} (4.12)

where width$_{i_k}$ is the width of compartment $C_{i_k}$ in high resolution pixels.

4.7 Generated Synthetic Data

In this section, two examples of synthetic data generated with the simulator are presented. The first is a 4D dataset of simulated kidney data, including movement artefacts and a tracer model derived from clinical DCE-MRI studies. This data has been used for the assessment of the 3D movement correction technique described in Section 3.4. The second example are various simulations of a 3D phantom that has been used to simulate the PV effect in small structures, using various noise conditions and PSFs.

4.7.1 Synthetic 4D DCE-MRI Kidney Data

The synthetic dataset consists of aorta and kidney objects, the latter consisting of cortical and medullary compartments. Multiple volumes were created, using a movement model, and a tissue and tracer functional model. The PSF model for DCE-MRI used was the Gaussian-Sinc, defined in Equation 3.9. Exemplar image slices can be seen in Figure 4.4.

The tissue and tracer functional models (following the approach of de Priester et al. (2003), described in Section 2.6.2) were derived from ROI data for both the cortical and medullary components. The variances in the ROIs were used to generate Gaussian distributed noise. As suggested in Section 4.5.2, noise in MR becomes Gaussian for large SNR. In the following section, the SNR on DCE-MRI data of the kidneys is estimated, and the Gaussian assumption is tested.
Figure 4.4: Simulated renal data. (Top) image slices containing the aorta and kidney from simulated data: at aortic-peak (left) and beginning (middle) and end (right) of the filtration phase. (Bottom) different image slices from a single volume at the cortical-peak.

Noise in the DCE-MRI Data

To test the validity of a Gaussian assumption for the noise in the magnitude data, the SNR of pre-contrast image data from the VIBE sequence proposed for clinical studies was estimated. A widely used method for estimating SNR is to define two ROIs (Firbank et al., 1999; Kaufman et al., 1989), placed on the tissue of interest and on background areas respectively. The first ROI, placed in the most homogeneous area of tissue is used to record the mean signal intensity $\mu_{ob}$. The second ROI is used to calculate the standard deviation $\sigma_{air}$ for the largest possible ROI placed outside the object in the image background. The SNR is then:

$$\text{SNR} = 0.655 \times \frac{\mu_{ob}}{\sigma_{air}}$$

where 0.655 is a correction factor to account for the effect of the magnitude operator on normally distributed (zero-mean Gaussian) noise (Henkelman, 1985).
4.7. Generated Synthetic Data

Repeated estimates of $\mu_{ob}$ and $\sigma_{air}$ were obtained from pre-contrast data volumes from 3 volunteer datasets. Average $\sigma_{air}$ and associated standard error were computed from two ROIs placed on background on two separate slices. Average $\mu_{ob}$ and associated standard error for both left and right kidneys were also collected (see Table 4.1).

Table 4.1: Mean signal intensity and noise variance in MR Renography data acquired using the VIBE sequence, using both right and left kidneys from three volunteers.

<table>
<thead>
<tr>
<th>ROI Intensity</th>
<th>$\sigma_{air}$</th>
<th>$\mu_{kidney}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>average</td>
<td>2.3</td>
<td>38.2</td>
</tr>
<tr>
<td>standard error</td>
<td>0.6</td>
<td>2.3</td>
</tr>
</tbody>
</table>

The resulting SNR value for pre-contrast kidney tissue was found to be $> 11$. After the arrival of tracer to these areas, SNR values may be considerably higher. Thus, this high SNR suggests that the noise distribution in MR Renography data using the proposed acquisition sequence may be approximated by Gaussian noise. An example of a Gaussian fit to data from one of the renal ROIs can be seen in 4.5, confirming this assumption.

Figure 4.5: Signal intensity distribution from 12 pre-contrast consecutive data volumes, placed on a right kidney ROI for one of the volunteers. The histogram data (points) provide a good fit to a Gaussian distribution (line).
4.7.2 Synthetic 3D PET Phantom Data

A synthetic phantom was used to investigate and assess the PV effect on small structures. It was based on a real PET phantom, consisting of a set of 3D simulated high intensity cylindrical inserts of different width on a low intensity background. The cylinders were angled at 10° to the horizontal and vertical planes for good sampling of the PV distribution. In the high resolution dataset representing an approximation to a continuous 'real world' cylinder, each voxel was assigned to class \( \omega_{bk} \) (background) or \( \omega_{ob} \) (cylinder) respectively. Thus, each voxel was assigned means of \( \mu_{bk} \) and \( \mu_{ob} \), and then Gaussian distributed noise components were added with respective standard deviations \( \sigma_{bk} \) and \( \sigma_{ob} \). By setting varying the standard deviation of the noise component, data volumes with different Contrast-to-Noise Ratio (CNR) were simulated.

CNR is a measure of class separability that takes into account not only the distance between the class means but also the noise level. It may be formally defined as:

\[
\text{CNR} = \frac{|\mu_{ob} - \mu_{bk}|}{\sigma} \tag{4.14}
\]

where \( \sigma \) is a measure of the noise in the system. In our simulations, it was defined as \( \sigma = \sigma_{bk} = \sigma_{ob} \).

Figure 4.6 presents image slices from phantom simulations with varying amount of noise and a Gaussian PSF corresponding to using \(^{18}\text{F}\) as a radioisotope on a PET scanner.

![Image slices from phantom simulations](image_url)

(a) No noise  (b) CNR = 50  (c) CNR = 5

Figure 4.6: Trans-axial views of simulated PET phantom data. The synthetic phantom consisted of a set of inserts of variable width and acquisition of the data by a PET scanner was simulated by a Gaussian PSF corresponding to using \(^{18}\text{F}\) as a radioisotope source.
4.8 Conclusions

In this chapter an overview of the synthetic data generator and some exemplar data generated have been presented. A generic approach has been implemented where the effect of the image acquisition system on high-resolution templates that represent an approximation to real world physical objects is defined by a user-defined PSF. This allows the generation of simulated data for distinct acquisition systems such as PET or MR sequences. The characteristics of noise, movement, SNR/CNR, and downsampling are also user-defined.

Synthetic data offers intrinsic GT, and therefore is useful in the assessment and development of image processing techniques, such as those for movement or PV correction. Data generated with the simulator has been used for these purposes in Chapters 3 and 6.
Chapter 5

Partial Volume Correction of Renograms

In this chapter, the importance of the Partial Volume (PV) effect in quantifying renal function from Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) data is explored. A methodology is developed that includes correction for movement of the kidneys due to respiration in the dynamic sequence. This is followed by the estimation of the tissue mixing for each voxel within a renal Region of Interest (ROI) from the convolution of the dynamic sequence's Point Spread Function (PSF) and high-resolution templates for each tissue type. Finally, the observed intensity for the ROI is de-composed, using the registered tissue mixing maps, into the true intensities that would correspond to pure voxels of each of the constituent tissues. Thus, non-renal contributions emanating from liver, spleen and other surrounding tissues can be eliminated from time-intensity curves derived from a typical renal cortical ROI.

The theoretical description of the PV correction step is presented in the next section. This is followed by a detailed description of the overall suggested methodology in Section 5.2. The results of applying this technique to a dataset of 10 healthy volunteers, acquired using the image acquisition protocol suggested in Chapter 3 are then presented and analysed in Sections 5.3 and 5.4.
5.1 Theory

Template-based PV correction approaches (see Section 2.5.1) are based on the assumption of having accurate (registered and PV un-affected) representations of the tissues or organ boundaries of interest. The effect of the PSF (of the dynamic sequence) on each of these *noiseless* tissue descriptors, which is described by their convolution with the PSF, represents the amount of blurring (both as overspill and underspill) exerted by the acquisition process on that tissue and, therefore, the extent of influence (and mixing) that it has on neighbouring voxels. Thus, these convolved templates are referred as PV mixing maps.

5.1.1 Construction of a PV Mixing Map

Mathematically, the problem of partial volumes might be thought of as finding, for every voxel $\omega$, a mixing vector, $\alpha_\omega$:

$$\alpha_\omega = \{\alpha_{\omega 1}, \alpha_{\omega 2}, \ldots, \alpha_{\omega n}\}$$

(5.1)

where $n$ is the number of tissue signals contributing to the overall observed signal intensity in a particular voxel. Thus, assuming linear mixing of the contributing MR signals within a particular voxel, the mixing vector represents the contribution of each tissue class $j$ to the observed intensity of a particular voxel $\omega$, so that:

$$\sum_{j=1}^{n} \alpha_{\omega j} = 1$$

(5.2)

One possible approach for determining $\alpha_\omega$ is through the convolution of binary tissue templates, which are an approximation to the real-world tissue boundaries, with the PSF of the imaging system. Thus, for a volumetric binary tissue template $T_j(\omega)$, its corresponding mixing map $A_j(\omega)$ is given by its convolution with the PSF:

$$A_j(\omega) = T_j(\omega) * \text{PSF}(\omega)$$

(5.3)

Therefore, $A_j(\omega)$ represents the contribution of tissue $j$ to every voxel $\omega$ within the volume.
5.1. Theory

Extraction of the Renal Component

Having obtained the mixing vectors, the observed intensity \( o_\omega(t) \) for each voxel at time \( t \), within or surrounding the kidney can be defined as:

\[
o_\omega(t) = \sum_{j=1}^{n} a_{j\omega} \times i_{j\omega}(t)
\]

where \( i_{j\omega} \) represents the true unmixed underlying intensity for tissue class \( j \), as measured at voxel \( \omega \).

Knowing \( o_\omega(t) \) and the mixing maps, the above equation can be solved to obtain the signal intensity curves of each individual tissue component. For a vector \( o(t) = [o_1(t), o_2(t), ..., o_\omega(t)]^T \), containing observed intensities from voxels within an ROI, the above equation can be written in matrix form as:

\[
o(t) = A i(t)
\]

where \( i(t) = [i_1(t), i_2(t), ..., i_\omega(t)]^T \) are the corresponding unmixed intensities and \( A \) is a \( \omega \times n \) matrix with elements \( a_{kj} \), that is the contribution from the \( j \)-th tissue on voxel \( k \).

For a typical ROI where there are more observations (voxels) than tissue types, Equation 5.5 defines an overdetermined system, with no inverse for \( A \) and no single solution. A possible solution may be a vector \( i' \) that minimises the Euclidean norm:

\[
\|Ai' - o\|^2
\]

Such a vector, \( i' \), can be obtained though the pseudo-inverse of \( A \), \( A^+ \):

\[
i' = A^+ o
\]

Thus, an optimal solution (in a least squares sense) for the component intensities at any time \( t \) may be obtained by calculating \( A^+ \). \( A^+ \) is a generalization of the inverse, and exists for any \( m \times n \) matrix. A common way of calculating \( A^+ \), as used in this work, is through singular value decomposition.
5.2 Methodology

In this section, a methodology for PV correction of MR renograms is presented. Figure 5.1 provides a schematic diagram of the flow of data processing used in the methodology.

Figure 5.1: Renal PV correction methodology.
The proposed approach is the result of several iterations and refinements that included an initial 2D analysis of the PV effect in the kidneys, using a simpler 2D Sinc model of the in-plane PSF. This preliminary work, which can be found in Rodriguez-Gutierrez et al. (2006), already suggested the importance of the PV effect on MR Renography data as the resulting PV-corrected renal curves presented an increased filtration slope, qualitatively closer to the expected renal curve than the uncorrected curve.

This preliminary approach was extended to include a more accurate 3D Gaussian-Sinc model of the PSF, as described in Chapter 3, and a quantitative assessment of the renal curves using Patlak analysis. The individual processing steps used in this work are discussed below.

5.2.1 Estimation of the Mixing Maps

The first step involved in constructing a PV mixing map was estimating the tissue-mixing produced by the PSF of the proposed dynamic data acquisition sequence. The high resolution templates were extracted from data captured using breath-holding techniques and the anatomical sequence described in Section 3.1. The anatomical data were manually segmented by a clinical expert into a set of 3D binary tissue templates, \( T_j \), including kidneys, liver, and spleen, as per the contours illustrated in Figure 5.2. An extra template was created, by default, from all voxels excluded from the afore-mentioned templates and regarded as background.

Each of the \( n \) tissue templates was then separately convolved with the 3D PSF, as per Equation 5.3, representing the intrinsic impulse response of the dynamic sequence (see Section 3.3.4) on each tissue. This produced a blurred representation of each functional tissue class corresponding to \( A_j \). The set of all mixing maps represents the action of the point spread function at each voxel and therefore:

\[
A(\omega) = \{A_1(\omega),...,A_n(\omega)\}
\]

is the matrix representation of the mixing vectors in the data volume as defined in Equation 5.1 Where the blurred templates overlap, the relative magnitude of each overlapping contribution describes the relative signal contribution from each adjacent
tissue class to the observed intensity of the voxel. Where there is no overlap in the templates, then the voxel may be considered to be a ‘pure’ voxel containing functional information from only a single tissue class. Otherwise, voxels were considered ‘mixed’ voxels:

\[ m(\omega) = \begin{cases} 
1 & \alpha_{\omega j} = 1 \quad \text{and} \quad \alpha_{\omega l} = 0 \quad \forall l \neq j \\
0 & \text{otherwise}
\end{cases} \quad (5.9) \]

The above classification would result in most voxels considered as mixed. To limit the effect of relatively small contributions a threshold was set to eliminate those below 1%:

\[ \alpha_{\omega j} = \begin{cases} 
\alpha_{\omega j} & \text{if} \quad \alpha_{\omega j} \geq 0.01 \\
0 & \text{otherwise}
\end{cases} \quad (5.10) \]

Therefore, the threshold eliminates negligible contributions and the map of pure voxels provides an indication of what areas of tissue are (mostly) free of other tissue contributions.
5.2. Methodology

5.2.2 Movement Correction and Template Registration

In order to accurately assign a valid mixing vector to each voxel in the dynamic sequence it was necessary to register the movement-corrected dynamic data and the mixing maps. The DCE-MRI data was movement corrected by the intensity-based registration in Denis de Senneville et al. (2006b), using the intensity correlation-based distance minimisation technique described in Section 2.3.3. It is based on an iterative gradient-descent optimization process, up to a step of 0.5 pixels and 0.5° respectively for translation and rotation parameters. In order to minimise the influence of intensity changes due to the passage of tracer, this movement correction is first applied to c pre-contrast data volumes \([V_1, ..., V_c]\) and the movement vectors associated with these, \([a_1, ..., a_c]\), are learned to produce a movement atlas for these pre-contrast data:

\[
\text{Hist} = \{\{V_1, a_1\}, ..., \{V_c, a_c\}\}
\]  

(5.11)

Then, subsequent (contrast-affected) search data volumes \(V_s\) are first matched against Hist (using the same correlation coefficient, see Equation 2.12), before their movement with respect to \(V_r\) is estimated. The movement vector associated with the closest volume pre-contrast volume is used to initialise the search between \(V_r\) and \(V_s\). This technique was used instead of the Fourier-based method implemented by the authors (based on translation-only movement correction), and described in Chapter 3, as it provided more accurate movement correction (based on visual inspection) of the dynamic data.

Having movement-corrected the dynamic data, it was then necessary to register it to the mixing maps. The high resolution data was acquired with the same scanner and within the same study, using the same FOV and choosing the anatomic data voxel size as integer multiples of the dynamic data. Therefore, errors associated with interpolation and re-binning were minimised in up-sampling the dynamic data to the same resolution of the anatomic data. Up-sampling was performed using bi-cubic interpolation.

Finally, the up-sampled dynamic and anatomic data (and therefore the mixing templates) were registered using the Fourier-based approach described in Section 3.4 as, with clearly distinct intensity distributions (see Figure 3.1), a non intensity-based approach was preferred.
5.2.3 Extraction of Tissue Component Curves

Having corrected any movement in the dynamic data, uncorrected mean observed intensity curves, \( \bar{x}(t) \), may be obtained for a given tissue ROI. For a renal ROI \( R \):

\[
\bar{x}(t) = \frac{1}{\Omega} \sum_{k=1}^{\Omega} o_k(t)
\]

(5.12)

where \( \Omega \) is the number of voxels in \( R \).

Such a curve may be de-composed into its constituent components by first obtaining the observed intensity vectors \( o_k \) for ROI \( R \) on every dynamic data volume at time \( t \).

Then, matrix \( A \) can be found from \( A \), where:

\[
A(k, j) = \alpha_{kj} \quad \forall k \in R
\]

(5.13)

PV-corrected curves for each tissue contributing to the ROI may then be found by Equation 5.5.

In this work, renal ROIs were only drawn on the centre-most slice of \( T_j \). The aim of this work was not to compute overall kidney GFR but to compare PV corrected and uncorrected GFR estimates. Therefore, using only 2D ROIs was aimed at reducing potential errors arising from the segmentation of the tissue templates in peripheral slices, as even on the high resolution data tissue boundaries on the outer-most were affected by PVs and therefore difficult to segment, even for expert clinicians.

5.2.4 Extraction of Arterial Input Functions

To calculate and compare GFR estimates, both PV corrected and uncorrected renal curves were used in Patlak analysis. In addition to the renal curves, an estimate of the AIF is also necessary for Patlak analysis. Mean AIF curves for an arterial ROI can be obtained in the same manner as the mean renal curve in Equation 5.12. Figure 5.3 shows a typical arterial ROI and its corresponding AIF.
Figure 5.3: Arterial ROI and corresponding Arterial Input Function. The ROI thickness perpendicular to the direction of flow is minimised to reduce time-averaging of the signal.

5.3 Results

In this section the effects of the PVE on MR Renography for a set of 10 volunteers is presented. Initially, the contributions from different tissues to typical ROIs from the centre-most renal template slice drawn on the cortex are presented. Using these PV estimates, the contributions from adjacent tissues are assumed to be eliminated from the the time-intensity curve to produce a PV-corrected estimate of the renal curve. To assess the robustness of the method a much larger abdominal ROI is then used in the same manner to extract PV component curves of the liver/spleen and compare these estimates with GT time-intensity curves extracted from single-tissue (PV un-affected) ROIs on the liver/spleen. Finally, the effects of the PV correction on the Patlak analysis of the data are presented.

5.3.1 Effect of Non-renal Contributions to Cortical ROIs

Figure 5.4 shows the contributions from different tissues to typical cortical ROIs for one of the volunteers, manually drawn by a clinical expert. The blurring exerted by the acquisition with the dynamic sequence produces a tissue-mixing within the ROI where $\approx 25\%$ of the contributions are due to non-renal components. The results presented in Figure 5.4 are consistent with those from the other volunteers, as can be seen in
Figure 5.4: ROI-overall and per-voxel tissue contributions. The contribution from each tissue, as a percentage of the ROI total for the right (left image) and left (right image) kidneys, for a cortical ROI is shown in the bar graphs. The top image show an overlay of the cortical ROIs and the per-pixel tissue mixing maps: these are colour-coded representation of the number of tissues contributing to each voxel where voxels with signal contributions from 1 tissue are in red, 2 tissues in green and 3 tissues in blue.

Table 5.1: Figure 5.4 also shows the number of components per voxel in the ROI.

Table 5.1: Component tissue contributions to cortical ROIs for each volunteer.

<table>
<thead>
<tr>
<th>volunteer</th>
<th>Right Kidney cortical ROI</th>
<th></th>
<th></th>
<th>Left Kidney cortical ROI</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kidney</td>
<td>Liver</td>
<td>Background</td>
<td>Kidney</td>
<td>Spleen</td>
<td>Background</td>
</tr>
<tr>
<td>1</td>
<td>80.1</td>
<td>0.8</td>
<td>19.1</td>
<td>75.3</td>
<td>0.3</td>
<td>24.5</td>
</tr>
<tr>
<td>2</td>
<td>69.7</td>
<td>3.8</td>
<td>26.5</td>
<td>73.6</td>
<td>0.0</td>
<td>26.4</td>
</tr>
<tr>
<td>3</td>
<td>65.7</td>
<td>1.3</td>
<td>33.0</td>
<td>64.7</td>
<td>0.4</td>
<td>34.9</td>
</tr>
<tr>
<td>4</td>
<td>82.9</td>
<td>0.7</td>
<td>16.4</td>
<td>81.2</td>
<td>0.0</td>
<td>18.8</td>
</tr>
<tr>
<td>5</td>
<td>76.3</td>
<td>1.1</td>
<td>22.6</td>
<td>67.8</td>
<td>2.4</td>
<td>29.8</td>
</tr>
<tr>
<td>6</td>
<td>75.6</td>
<td>0.9</td>
<td>23.5</td>
<td>76.0</td>
<td>0.1</td>
<td>24.0</td>
</tr>
<tr>
<td>7</td>
<td>79.3</td>
<td>1.1</td>
<td>19.6</td>
<td>78.3</td>
<td>0.7</td>
<td>21.0</td>
</tr>
<tr>
<td>8</td>
<td>72.1</td>
<td>3.2</td>
<td>24.7</td>
<td>72.7</td>
<td>0.7</td>
<td>26.6</td>
</tr>
<tr>
<td>9</td>
<td>81.0</td>
<td>1.5</td>
<td>17.5</td>
<td>78.0</td>
<td>2.8</td>
<td>19.2</td>
</tr>
<tr>
<td>10</td>
<td>76.3</td>
<td>4.2</td>
<td>19.5</td>
<td>78.5</td>
<td>1.3</td>
<td>20.2</td>
</tr>
<tr>
<td>Average</td>
<td>75.9</td>
<td>1.9</td>
<td>22.2</td>
<td>74.6</td>
<td>0.9</td>
<td>24.5</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>5.4</td>
<td>1.3</td>
<td>5.0</td>
<td>5.1</td>
<td>1.0</td>
<td>5.1</td>
</tr>
</tbody>
</table>
5.3. Results

5.3.2 Component curves

The cortical ROIs in the previous section were used to extract cortical time-intensity curves. The different component curves (renal, spleen, background, etc...) are estimated from these cortical ROI curves through a PV correction step. Cortical ROI curves and estimated renal component curves for one of the volunteers can be seen in Figure 5.5. The removal of non-renal components results in a marked increase of the curve in the filtration phase.

![Component curves](image)

**Figure 5.5:** Raw cortical ROI curves, \( I(t) \), (black lines) and estimated renal component curves (blue lines) for the right (left image) and left (right image) kidneys.

In the absence of GT for the renal curves, a possible approach to validate the proposed methodology is to compare another of the estimated components: the liver or the spleen component (for the right and left kidneys respectively). Both these organs are relatively large, and therefore can provide a suitable region of voxels unaffected by PV mixing, otherwise known as pure voxels, from which to extract pure liver and spleen GT curves. Further, they may be considered as to provide a relative homogeneous enhancement response to the contrast agent, unlike the distinctly heterogeneous response profiles that may be expected in the kidneys for cortical, medullary and collecting system tissues.

The PV contributions to the cortical ROIs above of both the liver and spleen are very small (see Figure 5.4) as most of the contributions are from the renal components. It proved impossible to reliably extract liver/spleen time-intensity curves from such small contributions to the overall signal. Therefore, to reliably estimate liver and spleen components, much larger ROIs, with considerable contributions from these tissues, are
used (see Figure 5.6). These rectangular-sized ROIs would not normally be drawn by a clinician for the purpose of extracting a renal time-intensity curve; however, these do allow a more accurate extraction of all the obfuscating tissue components, and help test the robustness of this approach.

![Image of ROIs](image)

**Figure 5.6:** (Top) Large abdominal ROIs used to extract renal and liver/spleen components for the right (left column) and left (right column) kidneys. (Bottom) Corresponding raw time intensity curve and extracted components.

Estimates from these large ROIs for the liver and spleen components can now be compared with pure liver and spleen GT curves, as in Figure 5.7. It can be observed that the recovered liver and spleen component curves closely coincide with their respective pure GT curves.

The renal component curves extracted from the large abdominal ROI and the cortical ROIs may also be compared to each other to evaluate the robustness of the proposed PV correction methodology in estimating the renal component curves, irrespective of the size and shape of the ROI used. It can be seen in Figure 5.7 how the method recovers the same renal components for both the large and the cortical ROI.
5.3. Results

Figure 5.7: (Top) Accuracy of component estimation: GT and liver and spleen curves. (Bottom) Robustness to ROIs size: comparison of renal components curves extracted from cortical and large abdominal ROIs.

5.3.3 Patlak Analysis

The effects of partial PV correction on the renal curves are presented by applying Patlak analysis (described in Section 2.1.1) to both the raw and corrected data. By using the temporal period between 140 and 200 sec, it can be assumed that negligible amount of contrast agent have been passed into the medullary region. It is therefore assumed that the cortical region behaves as a net accumulator of contrast agent, consistent with the assumption of Patlak analysis. The PV correction step results in an increase of the Patlak slope ($k_1$ in Equation 2.10) and a consistently better fit of the data as can be observed in Figure 5.8 and Table 5.2. The Patlak slope may be considered indicative of GFR and therefore with the PV correction step, results in a higher GFR estimate as the overspilling effects of neighbouring tissues are removed.
Figure 5.8: Patlak analysis of raw (Top) PV corrected data (bottom) for the right (left column) and the left (right column) kidneys.

Table 5.2 shows the effect of the PV correction stage on the Patlak analysis for all volunteers. There is a consistent increase in the slope, and therefore GFR, and the $R^2$ value of the fit following PV correction for every one of the analysed kidneys, thus demonstrating enhanced model fit.
5.3. Results

Table 5.2: Patlak results for raw (before correction) and PV corrected data showing, as a percentage, the increase in slope produced by PV correction. The mean percentage increase in $k_1$ for each kidney after PV correction across all volunteers was 36%.

<table>
<thead>
<tr>
<th>volunteer</th>
<th>left kidney before</th>
<th>left kidney after</th>
<th>%incr.</th>
<th>right kidney before</th>
<th>right kidney after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.012</td>
<td>0.018</td>
<td>35</td>
<td>0.92</td>
<td>0.95</td>
</tr>
<tr>
<td>2</td>
<td>0.015</td>
<td>0.023</td>
<td>36</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td>3</td>
<td>0.013</td>
<td>0.021</td>
<td>39</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td>4</td>
<td>0.019</td>
<td>0.029</td>
<td>33</td>
<td>0.83</td>
<td>0.89</td>
</tr>
<tr>
<td>5</td>
<td>0.009</td>
<td>0.018</td>
<td>48</td>
<td>0.81</td>
<td>0.94</td>
</tr>
<tr>
<td>6</td>
<td>0.019</td>
<td>0.034</td>
<td>44</td>
<td>0.74</td>
<td>0.88</td>
</tr>
<tr>
<td>7</td>
<td>0.021</td>
<td>0.030</td>
<td>32</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>8</td>
<td>0.012</td>
<td>0.019</td>
<td>21</td>
<td>0.88</td>
<td>0.91</td>
</tr>
<tr>
<td>9</td>
<td>0.015</td>
<td>0.019</td>
<td>21</td>
<td>0.88</td>
<td>0.91</td>
</tr>
<tr>
<td>10</td>
<td>0.015</td>
<td>0.023</td>
<td>34</td>
<td>0.90</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Finally, table 5.3 shows a comparison of the estimated DRF (before and after PV correction) for each volunteer. There are no substantial changes to DFR estimation by using a PV correction stage.

Table 5.3: Estimated DRF before and after PV correction.

<table>
<thead>
<tr>
<th>volunteer</th>
<th>left kidney before</th>
<th>left kidney after</th>
<th>right kidney before</th>
<th>right kidney after</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>52</td>
<td>48</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>50</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>50</td>
<td>48</td>
<td>52</td>
</tr>
<tr>
<td>4</td>
<td>56</td>
<td>44</td>
<td>60</td>
<td>40</td>
</tr>
<tr>
<td>5</td>
<td>56</td>
<td>44</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>6</td>
<td>51</td>
<td>49</td>
<td>55</td>
<td>45</td>
</tr>
<tr>
<td>7</td>
<td>51</td>
<td>49</td>
<td>52</td>
<td>48</td>
</tr>
<tr>
<td>8</td>
<td>52</td>
<td>48</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>47</td>
<td>53</td>
<td>46</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>52</td>
<td>48</td>
<td>51</td>
<td>49</td>
</tr>
</tbody>
</table>
5.4 Discussion

The proposed approach for PV correction of renal curves is based on the estimation of mixing vectors for every voxel, derived from the convolution of registered high-resolution anatomical volumetric binary tissue templates and the PSF of the dynamic acquisition sequence. These mixing vectors are then used to extract the individual component curves from time-intensity curves generated through ROI analysis of movement corrected renal DCE-MRI data. Finally, GFR estimates are obtained from these PV corrected renal curves.

It is important to note that GFR estimates as described here only represent the mean of a set of per-voxel estimates, and that the hematocrit has not been subtracted from them. As such, these estimates can not be used as single kidney GFR, and moreover, the total volume of functional renal tissue for either kidney has not been calculated. However, they may be used as an estimation of differential function, and for this hematocrit subtractions only represents a bias that is equal for both left and right kidneys.

There are three main results to be discussed from the work presented here:

Significance of PV effects in DCE-MRI data of the kidneys

PV effects from non-renal tissues on a typical ROI of the renal cortex, as estimated through the convolution of high resolution templates and the PSF of a typical DCE-MRI sequence, have been quantified for the first time. The PV effect has been shown to be significant, with $\approx 25\%$ (SD $\approx 5\%$) non renal contributions, observed in cortical ROIs.

The mixing-maps also show that there are few pure-renal voxels on a cortical ROI. Even those voxels that are considered pure-renal for the purposes of this work would in reality contain contributions from three very distinct components with different functional responses (cortical, medullary and collecting system). Thus, this works suggests that, assuming a significant PV effect is also applicable to other typical DCE-MRI sequences, it may not be possible to obtain pure cortical signals directly from small ROIs placed on the outer rim of the kidney. This is in disagreement with prior work.
5.4. Discussion

by Giele et al. (2002), whose approach to eliminate intra-parenchyma PV contributions is based on the assumption that pure (unmixed) cortical curves can be successfully extracted without explicit PV correction. The work presented here suggests that this is unlikely for the VIBE (Rodriguez Gutierrez et al., 2008; Sance et al., 2006) or other similar volumetric interpolated GE sequences used currently in MR Renography (Lee et al., 2003; Song et al., 2005) that capture 3D abdominal volumes in a relatively small time-frame (~2-3 s).

Correction of non-renal PV contributions

The proposed PV correction has been shown to accurately recover component curves for the liver and spleen. This was found to be consistent for every kidney analysed. The estimated renal components show a qualitative difference with respect to the raw curves where there is substantial change in the shape of the recovered parenchymal renal curve with a steeper slope in the critical period of filtration (60 - 120 seconds).

Quantitatively, the PV correction step results in an average increase of \( \approx 36\% \) (SD \( \approx 8\% \)) on the relative GFR with an improvement of around 5 percent in \( R^2 \) fitting to the Patlak model for the cohort studied here. This may be particularly significant given that Rutland-Patlak analysis (without PV correction) of MR Renography data has been reported to underestimate GFR (compared to absolute \(^{51}\)Cr-EDTA methods) (Annet et al., 2004). The introduction of a PV step has not resulted in substantial changes to estimates of differential function, as the Patlak slope is increased by a similar magnitude for both kidneys by the PV correction stage.

Robustness of the method to ROI size

A substantial source of variability may arise in MR Renography analysis depending on the size of the ROI used. The proposed method appears to be robust to changes in ROI, as observed in the recovered TI curves. This is also observable from the repeatability of Patlak slopes. As an example, Table 5.4 shows the effect of the PV correction on three different size ROIs for one volunteer. As can be seen, the Patlak slope estimates...
after PV correction are highly consistent, even considering the differences between the ROIs.

Table 5.4: Patlak results before and after PV correction for various ROI sizes for Volunteer 10. Patlak analysis was not carried out on data from the large ROI before PV correction as this ROI is not anatomically relevant.

<table>
<thead>
<tr>
<th>Left Kidney</th>
<th>Right Kidney</th>
</tr>
</thead>
<tbody>
<tr>
<td>k1</td>
<td>r2</td>
</tr>
<tr>
<td>before</td>
<td>after</td>
</tr>
<tr>
<td>small cortical</td>
<td>0.01</td>
</tr>
<tr>
<td>cortical</td>
<td>0.01</td>
</tr>
<tr>
<td>large abdominal</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 5.9: Small cortical ROI for Volunteer 10, used for comparing PV corrected Patlak results with those from the cortical and abdominal ROIs, as shown in Table 5.4.

The last ROI in Table 5.4 is particularly important. Generally, a clinician would not draw such an ROI for the purposes of estimating renal function. However, this work suggests that the same PV corrected GFR estimates may be obtained from this ROI as for the clinician-selected ROI. This suggests that the approach could be automated to routinely select an arbitrarily large abdominal ROI, eliminating the need for highly subjective expert ROI selection, and therefore reducing inter and intra observer variability in GFR estimation.
5.5 Conclusions

In this chapter, it has been shown that, for a cohort of 10 volunteers, within a region of interest naively thought to be pure cortex of the left and right kidneys, there was substantial contribution from the adjacent non-renal tissues (~25%). This is the first time that PV contributions from non-renal tissues to cortical ROIs have been investigated in MR Renography. Further, the importance of the PVE on the observed time intensity curves and subsequent Rutland-Patlak analysis for MR Renography has been demonstrated. An average increase in the Patlak slope of ~36% and an average improvement of 5% in the corresponding R² fitting was found for PV-corrected renal curves when compared to non-PV corrected ones. Whilst these increases in the Patlak slope do not result in significantly different DRF (both right and left kidneys are affected by PV effects of similar magnitude), they would have a substantial impact in absolute quantification of renal function. Given the consistency of these results, of both the enhancement from PV correction in k₁ and R² fitting in the Patlak analysis, then one might speculate that similar effects might be observed in other groups of volunteers and, moreover, when disease is present. Finally, the approach has been shown to be robust to changes in the size of the renal ROI.
Chapter 6

Statistical Partial Volume Correction of Small Structures

In the previous chapter, the Partial Volume (PV) effect on MR renograms was investigated. A methodology, based on the availability of high resolution organ templates, was proposed to extract PV-corrected renal curves. These were used, in conjunction with Arterial Input Functions (AIFs) also extracted from ROI analysis on the Dynamic Contrast Enhanced Magnetic Resonance Imaging (DCE-MRI) data, in estimating Glomerular Filtration Rate (GFR) through Patlak analysis. In that chapter, the PV effect on the AIF data was not considered.

In this chapter, the PV effect on small structures is investigated. Within the context of this work, a small structure is defined as having dimensions of similar magnitude as the PSF. In such cases, the PV effect may be an unavoidable corrupting influence affecting all voxels that capture the target structure. In cases where image intensities are extracted from ROIs placed on such structures for quantitative analysis, these intensities must be corrected. An example of such a structure may be a blood vessel over which an ROI is placed to estimate an AIF.

A template-based approach, similar to that of the previous chapter, could possibly be used for PV correction in the aorta. However, with the DCE-MRI and high resolution acquisition set-up optimised for capturing the kidneys, the aorta is not normally parallel.
Chapter 6. Statistical Partial Volume Correction of Small Structures

to the in-plane image slices, obfuscating segmentation of the high resolution templates. Further, due to the large voxel dimensions in the $z$-direction the aorta is only captured by two or three image slices and the PV effect, present in the high resolution data also, obfuscates the boundaries of the aorta. This effect, relevant to imaging of adults (with reported average abdominal aorta diameters ranging between 10 and 26 mm (Williams and Warwick, 1980), becomes more important in paediatrics, where the size of the blood vessel is smaller.

One alternative approach is to use a probabilistic PV classification technique such as the ones presented in Section 2.5.2. This avoids the need for registered templates and also mitigates the effects of template mismatch caused, for example, by dynamic respiratory movement. Based on Bayes theory, these techniques are used for labelling (classifying) voxels with a probability of class membership or mixing parameter. The limitations of such techniques when applied to small structures are discussed in Section 6.1. A method for addressing these problems by analytically modelling the mixing effects that PSFs have on an edge between two tissues is presented in Section 6.2. The approach is developed by using a Gaussian PSF, as found in many imaging systems. In the latter part of this chapter other PSFs are considered, particularly regarding those found in MRI imaging applications. The experimental set-up is described in Section 6.3 and the results are presented in Sections 6.4 and 6.5.

6.1 The PV Effect on Small Structures

In this section, the effect of filtering on small structures is investigated. The effects of 1D and 3D PSFs, using Gaussian PSFs as an example, on a finite-width step and cylinder respectively are considered analytically, to demonstrate that in situations where the width of the PSF is large in comparison to an object’s width, the PV effect may become so severe that the measured signal intensity from the object underestimates the true object intensity.
6.1.1 The Effect of a Gaussian Filter on 1D Finite-width Structures

Consider an imaging system with finite bandwidth and a corresponding spatially-invariant 1D PSF. In this case, the PSF is modelled as a Gaussian distribution $h(x, \sigma)$. Then, consider a noiseless idealised finite-width step $f(x)$ with unit height and width $w = k\sigma$, where $k$ is a constant (therefore, the width of the step is defined as a function of the width of the PSF, $\sigma$):

$$f(x) = \begin{cases} 1 & \text{for } 0 < x < k\sigma \\ 0 & \text{otherwise} \end{cases} \quad \text{(6.1)}$$

The result of convolving this finite-width step with $h(x, \sigma)$ is given by:

$$y(x, k, \sigma) = \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-\tau)^2}{2\sigma^2}\right) d\tau - \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma^2} \exp\left(-\frac{(x-\tau-k\sigma)^2}{2\sigma^2}\right) d\tau. \quad \text{(6.2)}$$

The first integral represents the accumulation of the acquired signal as $h(x, \sigma)$ approaches and passes over an object's rising edge in the convolution process associated with band-limited image acquisition. The second component represents a corresponding loss of signal as $h(x, \sigma)$ approaches and crosses the falling step edge of the object. A graphical representation of the process, described by Equation 6.2, can be seen in Figure 6.1.

![Figure 6.1: Effects of a Gaussian PSF (on an arbitrary scale) on a finite width step for which $f(x) = 1$ for $[0, k\sigma]$. The result is the difference between two error functions, see text for details.](image-url)
Equation 6.2 can also be expressed as:

\[ y(x, k, \sigma) = \frac{1}{2} \left[ \text{erf} \left( \frac{x}{\sqrt{2} \sigma^2} \right) - \text{erf} \left( \frac{x - k, \sigma}{\sqrt{2} \sigma^2} \right) \right], \tag{6.3} \]

where \( \text{erf}(p) = \frac{2}{\sqrt{\pi}} \int_0^p \exp(-r^2)dr \) is the Gaussian error function.

For objects of large width, \( w > 5.\sigma \), the second error function, centred at \( k\sigma \), has no effect on the rising edge (first error function) and \( y(x) \) reaches unit height. However, for small \( w \), the second error function reduces the height of \( y(x) \), which no longer reaches the original unit height of \( f(x) \). The results can be seen in Figure 6.2. For \( w < 5.\sigma \) the unit height is no longer attained. Further, the resulting function no longer resembles a rising error function followed by a plateau and a descending error function; instead it resembles a Gaussian-like function.

\[ \begin{array}{c}
\text{Figure 6.2: Effects of the convolution of a Gaussian PSF and a finite width step as a function of step width } k.\sigma. \text{ As can be seen, for smaller values of } k.\sigma, \text{ the curves no longer reach the true height of the original unit step.}
\end{array} \]
6.1.2 The Effect of a Gaussian Filter on 3D Finite-width Cylindrical Structures

The problem of true intensity loss due to the PV effect has been addressed in PET imaging by the estimation of Recovery Coefficients (RCs) to recover 'true' intensities. These can be obtained empirically: for example, Asselin et al. (2004) used a cylindrical model fitted to blood vessel data to calculate RCs and therefore estimate true intensities from observed intensities.

Another approach is, again, to model the imaged object as a geometric 3D shape and then derive analytically the effects on the shape of the PSF of the acquisition system. Thus, this is an extrapolation of the 1D method above. The description of the signal \( s(x) \) corresponding to the 3D PSF-blurring for a cylindrical object \( y \) of uniform intensity \( A \) and radius \( r \), is given by:

\[
s(x, r) = A \times y(x, r),
\]

where \( y(x, r) \) is the result of the convolution of the 3D PSF and the cylinder.

The effects of the 3D convolution for small cylinders are similar to those of the 1D case. That is, the maxima of \( y(r, x) \), \( \mathcal{L} \), may be smaller than one and therefore the maximum of Equation 6.4 , \( \max|s(x, r)| \), may be smaller than the true value, \( A \). Assuming that the cylinder's length \( L \) is much greater than its radius, and for, say, a Gaussian filter used to model the PSF, \( \mathcal{L} \) was derived by Kessler et al. (1984):

\[
\mathcal{L} = y(0, r/\sigma) = \left[1 - \exp\left(-x^2/2\sigma^2\right)\right] \times \left[1 - 2 \cdot \text{erfc}(L/(2\sqrt{2}\sigma))\right],
\]

where \( \text{erfc}(x) = 1 - \text{erf}(x) \).

The recovery coefficient \( \mathcal{L} \) was used by Kessler et al. (1984) for PV affected small cylinders captured with PET, and has been widely adopted in the PET imaging community. The difference in recovery coefficient, \( \mathcal{L} \), calculation from this 3D cylinder formulation and the 1D model from Equation 6.3 can be seen in Figure 6.3. The 1D model overestimates the convolution of the cylinder and 3D PSF.

Having presented the PV effect on small structures and how it affects the observed
maximum intensity of small objects, the next section introduces a PV classification and intensity recovery approach for small structures.

### 6.2 PV Classification of Small Structures

The PV classifier suggested by Wells et al. (2007) uses an explicitly defined mixing prior for a two class problem of background, $\omega_{\text{bk}}$, and object, $\omega_{\text{ob}}$. It assumes that two homogeneous tissues corresponding to these classes may be modelled by intensity distributions with mean intensities $\mu_{\text{bk}}$ and $\mu_{\text{ob}}$, and standard deviations $\sigma_{\text{bk}}$ and $\sigma_{\text{ob}}$. Using this approach, a mixing, $\alpha$, between two voxel classes, given an observed pixel intensity $x$, is given by:

$$E[\alpha|p(x|\alpha)] = \int_{\alpha} \alpha \frac{p(x|\alpha)p(\alpha)}{p(x)} d\alpha$$

(6.6)
where \( p(x|\alpha) \) is the likelihood of obtaining a particular signal value for a mixing value \( \alpha \), \( p(\alpha) \) is the \textit{a priori} Probability Density Function (PDF), or mixing density, reflecting the intensity mixing produced by the PSF of a particular imaging system, and \( p(x) \) is the marginal density for the measured signal.

The main advantages of this description is that the mixing density is defined explicitly, and that parameter estimation is simplified compared with other approaches in the literature (Laidlaw et al., 1998; van Leemput et al., 2003). The success of any of these classifiers relies on accurate parameter estimation of the class means \( \mu_{ob} \) and \( \mu_{ob} \).

However, as seen in the previous section, in the case of small structures, estimates of the object's mean intensity from image data, \( \tilde{\mu}_{ob} \), may be incorrect as the true mean \( \mu_{ob} \) may be lost due to the convolving action of the PSF. Therefore, for small structures there is a need for an intensity recovery step in addition to classification.

### 6.2.1 A Mixing Prior for Small Structures

The PVE on small structures not only affects the observed maximum intensity in the target object, but also influences the intensity distribution within (and immediately outside) the object. The mixing priors that have been used in the literature (see Section 2.5.2), such as the Benford or ICG models (Chiverton and Wells, June 2006), are all symmetrical functions that do not take the object's size into account and therefore fail to model correctly the asymmetric mixing distribution that occurs for small structures. The natural extension to the ICG prior to model correctly the 3D PV effect on small cylindrical structures would require the inversion of Equation 6.4. However, this equation is by definition, as any multi-dimensional function, non-invertible.

An alternative approach is to still use a 1D model, similar to ICG, that is suitable for small cylinders. One such model is the inverse of a 1D Gaussian. Noting its symmetry, and following the approach of Chiverton and Wells (June 2006) for the ICG, the rising edge of a 1D Gaussian, i.e. \( x \in (-\infty, \mu] \), can also be inverted and we refer to its gradient as the Inverse Gaussian (IG) PDF:

\[
p_{IG}(\alpha) = \begin{cases} 
C_{IG} \times \frac{\sigma}{\alpha \times \sqrt{-2 \ln(\alpha)}} & \text{for } \alpha \in (0, L] \\
0 & \text{for } \alpha \in (L, 1)
\end{cases}
\]  

(6.7)
where \( C_{IG} \) is a normalization constant and \( L \) is the recovery coefficient. For a cylindrical object, \( L \) depends on the diameter of the cylinder and is given by Equation 6.5. Thus the recovery coefficient is used to control the mixing probability range. The effect that this has can be seen in Figure 6.4, where the asymmetric IG curve (solid line) is now cropped for cylinders with recovery coefficient below one.

\[
\text{Figure 6.4: } p(\alpha) \text{ for various recovery coefficients. Small cylinders may be blurred by the PSF such that their 'true' intensity is no longer observable. In such cases, the mixing probability, } p(\alpha), \text{ truncates to zero for } \alpha > L.
\]

### 6.2.2 Intensity Recovery

The classifier in Equation 6.6 determines the mixing, \( \alpha \), between two voxel classes. Considering the estimates from image data of background and object means \( \bar{\mu}_{bk} < \bar{\mu}_{ob} \), the PV-corrected intensity for a voxel \( \omega \) may be found, as described in Equation 6.8, by:

\[
\sigma_\omega = (1 - \alpha) \times \bar{\mu}_{bk} + \alpha \times \bar{\mu}_{ob} \tag{6.8}
\]

However, it has been shown that for small structures, the average signal intensity for an object \( \mu_{ob} \), may not be observable and its estimate, \( \bar{\mu}_{ob} \), would underestimate the
true value. In such cases, a closer estimate for $\mu_{ob}$ than $\tilde{\mu}_{ob}$, $\tilde{M}_{ob}$, might be obtained by:

$$\tilde{M}_{ob} = \tilde{\mu}_{bk} + \frac{\tilde{\mu}_{ob} - \tilde{\mu}_{bk}}{L}$$

(6.9)

where $L$ is the recovery coefficient, as per Equation 6.5. Thus, substituting $\tilde{M}_{ob}$ for $\tilde{\mu}_{ob}$ in Equation 6.8, an intensity correction step for small structures is added to the PV correction.

6.3 Simulated data

We can illustrate the PV effect on small structures through simulation of a phantom containing cylinders of different diameters. The availability of ground truth for these simulated data enables the true a priori PV distribution to be observed. A similar simulated phantom to that described in Section 4.7.2 was produced. It consisted of a set of high intensity cylinders of different sizes on a low intensity background. The sizes of the cylinders ranged from $k = 1$ to $k = 9$ where $k = 2r/\sigma_{PSF}$. Thus, the data were filtered with a 3D Gaussian kernel, and down-sampled using decimation to a lower resolution grid. By setting $\sigma_{bk} = \sigma_{ob}$ and varying the object's mean intensity $\mu_{ob}$, data volumes with different CNR were simulated.

Simulated ground-truth, or an idealized representation of the image data in the absence of any corruption artefacts, was produced by ignoring any form of additive noise contribution before the filtering and down-sampling steps. The ground truth shows the intrinsic mixing produced by the PSF, without contamination from noise: it is this distribution that the mixing prior describes.
6.4 Results and Discussion

In this section, results are presented for PV classification using the IG model for the simulated cylindrical data. First, the IG model is compared with the ICG in terms of modelling the histogram of small cylindrical data. The results of the PV classifier using the IG prior for cylinders of different widths as a function of CNR are then presented for IG and several other priors. Finally, results for the intensity recovery step are presented.

6.4.1 Inverse Gaussian Model and Cylindrical Structures

An exemplar histogram of the GT simulated data and its fitting with the proposed prior mixing Inverse Gaussian density shown in Figure 6.5. It illustrates excellent fitting, with the IG model accurately describing the histogram of even a relatively large cylinder \((k=8)\). In this case pure voxels, and therefore \(\rho_{ob}\), are observable due to the large cylinder size. However, the PV effect produces an asymmetry present in the histogram related to the cylindrical shape. As the size of the cylinder increases further, and the pure-background to pure-object voxel ratio decreases, the asymmetry becomes less marked.

Figure 6.5: Inverse Gaussian Fit to histogram data of a small 3D cylinder \((k = 8)\) filtered with a Gaussian PSF. Note excellent fit between model and data, reflecting the expected asymmetry: the gentle drop between \((100,300)\) is followed by a much sharper rate of increase between \((300,450)\).
6.4. Results and Discussion

A comparison of the fitting of the IG prior to a cylindrical model where $\mu_{ob}$ is not observable (a smaller cylinder) can be seen in Figure 6.6. For comparison, the fitting of other mixing priors are also shown. The ICG Prior is defined in Equation 2.24, and the Truncated ICG Prior is set to zero in $(L,1)$, therefore, using the knowledge of the object radius, as in the IG case. As can be seen, the IG provides a closer fit than the other models for this cylinder width ($k = 4$), not only modelling the cylinder's asymmetry but also the lack of pure voxels.

Figure 6.6: Inverse Gaussian fit to histogram data of a small 3D cylinder filtered with a Gaussian PSF. The IG Prior is (purple) provides a better fit to the simulated data (squares) than either the ICG Prior (red) or the Truncated ICG Prior (blue).

6.4.2 PV Classification Results

In order to assess PV classification, a root mean squared (RMS) error measure is defined for every voxel, $v$ within ROI volume $V$ such that:

$$\text{RMS} = \sqrt{\frac{1}{V} \sum_{v} (\mu_{v} - \bar{\mu}_{v})^2}$$

where $\mu_{v}$ and $\bar{\mu}_{v}$ are the ground truth and classifier estimates for voxel $v$. 

$$(6.10)$$
Classification results for various cylinder widths are presented in Figures 6.7 and Figures 6.8 as the mean rms PV voxel error between the classified data and corresponding GT, for Uniform, ICG, IG, and a Histogram Prior. The latter was constructed from the histogram of the simulated data under no noise conditions. Therefore it represents the true mixing exerted by the PSF on the cylindrical objects.

Figure 6.7: Variation of classification error as a function of CNR for PV voxels on cylinders of several widths using uniform (top) and ICG (bottom) priors. The red curves denote 1/CNR.
Figure 6.8: Variation of classification error as a function of CNR for PV voxels on cylinders of several widths using IG (top) and Histogram (bottom) priors. The red curves denote $1/$CNR.
It can be clearly seen from these graphs that the results are very similar irrespective of the prior used. There is very little difference in these results as a function of cylinder width. For the smallest cylinders, below $k = 4$ (not plotted), a decrease of performance is experienced, with classification errors considerably higher.

It is also important to note the shape of all these curves: The increase in rms error as CNR decreases approximately follows a $1/\text{CNR}$ curve. These observations are consistent to those in Chiverton (2006), using several other priors for PV correction of larger structures, where $\mu_{ob}$ is easily observed.

As can be seen, the differences in classification performance between the different priors is small. In fact, the classification results seem to be slightly better for the Uniform prior, even than for the histogram prior, which was derived from actual histogram data of the noiseless cylinder simulation. This difference may be explained by the discretisation and re-binning steps that are performed on the priors before their use in the classifier. The Uniform prior is possibly less affected by these steps and the result is a slightly smaller error for those priors.

### 6.4.3 Intensity Recovery Results

Intensity recovery was performed on the classified data using analytically derived recovery coefficients. The recovered intensities were obtained from applying the recovery coefficients to $\tilde{\mu}_{ob}$ using Equation 6.9. $\tilde{\mu}_{ob}$ and $\tilde{\mu}_{ob}$ were estimated from GT data.

Table 6.1 shows $\tilde{\mu}_{ob}$ and $\tilde{M}_{ob}$ for simulated cylinders.

### 6.5 Other PSFs

In the previous sections, a classifier for PV correction has been proposed. The approach includes the use of analytically derived mixing priors. An alternative approach to analytically calculating recovery coefficients is to estimate these through simulation. To illustrate this approach, simulated data was generated using a 3D Sinc PSF and the recovery coefficient obtained from the average maximum intensity of noiseless histogram data for each simulated cylinders size. The resulting recovery coefficient curve may be
Table 6.1: Intensity recovery error for $\mu_{ob}$ as a percentage of the difference in means for different CNRs. The left-most column corresponds to a cylinder that is large enough for pure voxels to be observable.

<table>
<thead>
<tr>
<th>CNR</th>
<th>Estimate</th>
<th>$k_1$</th>
<th>$k_2$</th>
<th>$k_3$</th>
<th>$k_4$</th>
<th>$k_5$</th>
<th>$k_6$</th>
<th>$k_7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>$\hat{\mu}_{ob}$</td>
<td>89</td>
<td>62</td>
<td>33</td>
<td>14</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>59</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>$\hat{\mu}_{ob}$</td>
<td>89</td>
<td>61</td>
<td>33</td>
<td>14</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>$\hat{\mu}_{ob}$</td>
<td>89</td>
<td>62</td>
<td>33</td>
<td>14</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>10</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>$\hat{\mu}_{ob}$</td>
<td>89</td>
<td>62</td>
<td>33</td>
<td>14</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>$\hat{\mu}_{ob}$</td>
<td>89</td>
<td>63</td>
<td>34</td>
<td>15</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>$\hat{\mu}_{ob}$</td>
<td>90</td>
<td>62</td>
<td>33</td>
<td>15</td>
<td>5</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>17</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>$\hat{\mu}_{ob}$</td>
<td>90</td>
<td>62</td>
<td>34</td>
<td>16</td>
<td>6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>18</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>$\hat{\mu}_{ob}$</td>
<td>90</td>
<td>63</td>
<td>35</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$\hat{M}_{ob}$</td>
<td>18</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

seen in Figure 6.9. The recovery coefficient for the PSF of the DCE-MRI sequence, also generated through simulated data, may also be seen in the figure. For these non-Gaussian cases, $k$ was defined as $k = 4.7r/\text{FWHM}_{\text{PSF}}$.

6.5.1 Classification Results

The classifier, using the IG mixing prior, was applied to simulated cylinder data, following the same process as per the Gaussian case, for both the Sinc and DCE-MRI PSFs. Results for cylinders of width $k = 9, 6,$ and 4 are presented in Fig. 6.10 as the mean rms PV voxel error between the classified data and corresponding GT.

As in the previous cases that used a Gaussian PSF, classification results closely follow a 1/CNR trend, both for the Sinc and the Sinc-Gaussian PSF. Again, performance deteriorates for smaller values of $k (< 4)$. 
Figure 6.9: Recovery coefficient different size cylinders filtered with a Sinc PSF (crosses), the dotted line is to guide the eye only. The recovery coefficient for the Sinc-Gaussian PSF of the DCE-MRI sequence can also be seen (squares). The corresponding Gaussian recovery coefficient is also shown for comparison (solid line).

6.6 Conclusions

A method has been described that uses a statistical model to correct for partial volumes and in small cylindrical objects, corresponding to arterial vasculature. The method does not require expert ROI definition nor a step as in the template-based approaches for PV correction. This effectively means that this approach is immune to motion issues, provided the background statistics are approximately stationary.

PV classification results for several prior mixing models present a deterioration inversely proportional to CNR from $5 \rightarrow 25\%$ corresponding to CNR $20 \rightarrow 4$. There is no significant difference in the classification as a function of cylinder size for cylinders of width $k > 4$, either depending on the mixing prior used for classification, suggesting that this type of performance may be innate to these types of 1D intensity classifiers.

This independence on the mixing prior of classifier performance suggests that these classifiers may be used, without a detailed modelling of the scanner's PSF, to correct
Figure 6.10: Variation of classification error as a function of CNR for PV voxels on cylinders of several widths using the IG prior. The top image shows the results on simulated data generated using a Sinc PSF. The bottom image shows the results on data generated using the DCE-MRI PSF. The red curves denote 1/CNR.

PV effects in MR Renography data.

In addition to the classifier, a further intensity recovery step is suggested to be used, in order to recover voxel intensities blurred by the PSF. The recovery step is based on the calculation of a recovery coefficient, which has been investigated, analytically,
for a Gaussian PSF and, through experimental simulation, for a Sinc PSF and for the Sinc-Gaussian PSF of the DCE-MRI sequence.

The Sinc-Gaussian PSF of the DCE-MRI sequence presents an out-of-plane FWHM of 14.8 mm. This is approximately the mean size for aorta diameters reported in Williams and Warwick (1980) in adults, and this represents a $\approx 60\%$ intensity reduction according to the Sinc-Gaussian recovery coefficient from Figure 6.9. This effect could become more severe for paediatric patients which present smaller aortic diameters.
Chapter 7

Conclusions and Further Work

7.1 Summary of Results and Conclusions

The work presented in this thesis was concerned with image processing and analysis techniques for renography techniques using DCE-MRI data. The first chapters present an introduction to the kidneys and renal assessment, particularly using NI to assess differential function. These techniques form the basis of current research towards the use of MR Renography as a replacement for Radionuclide Renography due to, amongst other factors, increased spatial resolution.

Whereas lower resolution and absence of anatomical information means movement artefacts are not considered in Radionuclide Renography, movement correction is generally the first step taken in the analysis of DCE-MRI data of the kidneys. A method for 3D movement correction based on the phase difference of time-adjacent volumes was proposed in Chapter 3. Results on clinical data showed that translational movement was correctly estimated for 63% of the cases, with 80% of the movement correctly estimated to within 1 voxel when compared to manually generated GT.

In the absence of GT, similar work by Song et al. (2005) tested the movement correction results against synthetic-generated movement where a volume of data at time $t$ was arbitrarily shifted and rotated. These data therefore represents an oversimplification of the DCE-MRI problem as enhancement is a function of time, representing
Chapter 7. Conclusions and Further Work

not a static but rapidly changing spatial intensity distribution. In order to test the performance of movement correction under changing intensity conditions, a synthetic data generator was designed where the different kidney compartments may be assigned distinct enhancement profiles and 3D rotation and sub-voxel translation can be applied to the data. This was described in Chapter 4.

In addition to providing a tool for assessment of movement correction, and potentially of compartmental analysis such as Rutland-Patlak through ROI analysis, the main aim of the data simulator was to generate synthetic PV affected data to be used as GT to test PV correction approaches. Therefore, the most important part of the generator was concerned with applying the PSF of the simulated imaging system to a synthetic representation of the real world data. The generator was designed to use arbitrary PSFs, such as Gaussian kernels (as found in PET scanners) or Sinc-like kernels (typical DCE-MRI sequence). In the latter case, a thorough experimental analysis of the 3D PSF through phantom data was performed, and an empirical model for the PSF was proposed (Chapter 3).

PV correction in small structures, in particular cylindrical objects corresponding to arterial vasculature, with respect to the PSF were investigated in Chapter 6. A novel mixing prior, the IG which is an approximation to the convolution of a Gaussian and a finite width step was shown to model more closely noiseless histogram data than others proposed in the literature for small step widths. This, however, did not translate into a superior classification performance when applied to a PV bayesian classifier. Classifier performance was found to be approximately inversely proportional to CNR producing rms error of 5 → 25% corresponding to CNR 20 → 4, independent of cylinder size (for \( k > 4 \)) or mixing prior.

In addition to the classifier, a further intensity recovery step was suggested in order to recover the mixing components skewed by loss of the true mean intensity from histograms of small objects (dimensions \( \lesssim \) PSF). The recovery step is based on the calculation of a recovery coefficient, developed analytically for a Gaussian PSF and then, through experimental simulation, for Sinc-based PSFs. Results with synthetic data have shown that assuming accurate parameter estimates can be obtained, this
recovery can successfully estimate the original intensities blurred by the PSF for very small objects. The Sinc-Gaussian PSF of the DCE-MRI sequence presents an out-of-plane FWHM of 14.8 mm. This is approximately the mean size for aorta diameters in adults, and this represents a $\approx 60\%$ intensity reduction according to the Sinc-Gaussian recovery coefficient. Whilst this effect would be reduced for patients with larger aortas or improved sequences with higher resolution, it could also become more severe for paediatric patients which present smaller aortic diameters.

In Chapter 5, a method for quantifying and correcting PV contributions from extrarenal components in renal curves derived from dynamic Gd enhanced MR Renography was presented. For the first time, the contribution of external contributions to typical cortical regions of interest was investigated and reported to be in the region of 25%. A method for correcting these PV contributions was presented. Extraction of component curves was demonstrated by comparing the time-intensity plots for kidney and spleen components to pure (non PV affected) curves from ROIs drawn in the centre of these organs. In addition, the method proved to be largely independent of ROI size, potentially providing the route to an automated approach which is free from subjective errors in target ROI definition. Quantitatively, the effect of PV correction produced a mean increase of $\approx 36\%$ (SD $\approx 8\%$) on GFR estimates as calculated through Patlak analysis, with an improvement of 5% in the corresponding $R^2$ fittings.

This work has shown that in addition to movement correction, contrast agent quantification and accurate modelling, absolute GFR measurement via DCE-MRI is significantly affected by the PV Effect, and as such must be robustly accounted for.

### 7.2 Further Work

The GFR estimates presented in Chapter 5 can only be used for estimation of differential function. These represent mean ‘per voxel’ estimates and therefore need to be multiplied by the volume of functioning tissue in each kidney, that is the renal cortex, in order to produce single kidney GFR measurements. In order to achieve the, so far, unrealised goal of absolute GFR quantification, several further steps are needed. These include segmentation of the renal cortex and PV correction of the AIF and medullary
components on the cortex, and accurate quantification of the relationship between Gd concentration and signal intensity.

3D segmentation of the renal cortex would be the obvious next step required. Most information to differentiate between cortex and medulla is to be found in the time domain, corresponding to two different enhancement profiles and therefore, it would seem likely that methods based on clustering of voxels exhibiting a similar enhancement curve would produce good results. Any other method, relying purely on spatial information would most probably result in errors, particularly on the outer-most slices, due to severe PV effect affecting other than the centre-most slides. Although others (Zöllner et al., 2007; Sun et al., 2002) have already attempted to classify voxels based on their temporal functional behaviour, the work in this thesis demonstrates that such an approach could only be successful if combined with a voxel mixing or PV correction-based approach. Such an approach has not been published.

Regardless of the approach used for intra-kidney segmentation, the resulting regions would undoubtedly be affected by intra-kidney PV effect. Whilst the renal curves produced in this work have been corrected for extra-renal components, these still represent composite curves of three characteristically distinct functional components: cortical, medullary, and collecting system curves. However, it could be argued that pelvic and medullary components do not spill-over and add to the observed signal in the filtration stage. Nonetheless, their mixing presence needs to be rigorously accounted for as they will 'wash-out' some of the cortical component required for absolute quantification.

Therefore, more accurate GFR estimates could be expected after the removal of both medullary and collecting system components from the renal curve. The main difficulty in applying a template approach as for the external PV components lies in the absence of cortical and medullary boundary information on high resolution anatomical data. Any approach that used as templates the boundaries extracted from segmentation of the functional data, as suggested in the above paragraph, would be biased by the PV effect in the template segmentation step. An alternative approach could be the use of a statistical PV classifier, as per the one in Chapter 6, but using a simple uniform prior. Thus, the classifier would not need a priori knowledge of the size or shape of the intra-
7.2. Further Work

kidney regions which obfuscates the template definition problem. However, providing unique solutions to the resulting complex renal histogram would be non-trivial. A possible approach could be through the segmentation of a pelvical template from the high resolution data. Therefore, if this component could be eliminated in the same manner as other non-renal tissues (liver, background, etc.) the resulting parenchymal template would only consist of two tissue classes: renal cortex and medulla, simplifying the statistical classification process.

In this work, image processing techniques including a novel PV correction step have been applied to the MR Renography data in order to produce more accurate and robust results. However, in addition to incremental refinement of the current renography paradigm, MR offers the potential to develop novel approaches to renal function assessment. Whereas technological advances may reduce the effects of PVs on the whole-kidney GFR estimation presented here, the possible new applications that these advances would make possible, such as localised RBF or GFR estimation, would still be limited by the PV effect.
Appendix A

Anatomy of the Nephron and Renal Vasculature

A.1 The Nephron

The nephrons are the functional units of the kidney. There are about 1 million nephrons per kidney, and they produce urine through the processes of blood filtration, and the subsequent processes of secretion and reabsorption of water and other substances from the filtrate. Each nephron consists of a renal corpuscle and a renal tubule and is approximately 50 mm long (Gabriel, 1988). Blood is filtrated as it passes through the corpuscle to form a plasma ultra-filtrate whose composition is then further modified along the tubule by secretion and reabsorption of water and solutes back to and from the ultra-filtrate to the medullary interstitial tissue and blood plasma.

From the corpuscle, the ultra-filtrate passes to the proximal convoluted tubule, still located in the cortex. It then penetrates into the medulla via the proximal straight tubule and the thin descending limb, up to the loop of Henle, where it turns back towards the cortex via the thin and thick ascending limbs. Back in the cortex, the ultra-filtrate passes through the distal convoluted tubule and the connecting tubule until reaching the collecting ducts. Collecting ducts originate from merging nephronal tubules to form larger ducts deep into the medulla, known as ducts of Bellini, which directly
open into the surface of the renal papillae, from where urine finally flows into the drainage system of the kidney. Schematic colour-coded diagrams of nephron anatomy, depending on their location within the parenchyma, and of the renal vasculature that supplies them, can be seen in Figure A.1.

Figure A.1: Cross section diagram of the renal parenchyma: renal capsule (1), cortex (2), outer stripe of the outer medulla (3a), inner stripe of the outer medulla (3b) and inner medulla (4). The parenchymal vascular system, with the efferent arteriole (5), afferent arteriole (6), stellate vein (7), interlobular vein (8), interlobular artery (9), peritubular capillary (10), juxtamedullary efferent arteriole (11), arcuate artery (12), arcuate vein (13) and vasa recta (14). From Universität de Fribourg (2005)

Nephrons are normally characterised according to the location of their glomerulus within the cortex as to superficial, mid-cortical or juxtamedullary. Superficial nephrons have their glomerulus in the outer cortex and their loop of Henle extend only into the outer medulla. Mid-cortical nephrons have the glomerulus in the mid cortex and whilst some have short loops that reach into the outer medulla, most have loops of intermediate length that reach into the inner medulla. Finally, juxtamedullary nephrons have glomeruli in the inner cortex and long loops that reach into the tip of the papilla.
A.2 Renal Vasculature

In addition to supplying blood to the nephrons for filtration, the complex vascular network in the renal parenchyma has an additional important role in the creation of an osmotic pressure gradient, fundamental to the secretion and reabsorption processes. A diagram of the renal vasculature can be seen in Figure A.1.

The renal arteries originate in the aorta and once inside the kidney divide into several interlobar arteries. These further divide at the corticomedullary border into arcuate arteries, which arch over the base of the pyramids. From these, the interlobular arteries enter the cortex and divide into several afferent arterioles, supplying blood renal corpuscles through glomerular capillaries. After approximately 20% of blood plasma is filtrated, the remaining unfiltered blood exits the glomeruli via the efferent arterioles. Efferent arterioles form different structures, depending on the location of the glomerulus they supply: on one hand, efferent arterioles from mid-cortical and superficial nephrons form peritubular capillaries that supply blood to nephronal tubular segments within the cortex. On the other hand, efferent arterioles from juxtamedullary nephrons give rise to the descending vasa recta which penetrate deep into the medulla.

Blood is drained from the medulla by the ascending vasa recta, which merge with stellate and interlobular veins draining the cortex to form arcuate veins, which in turn form interlobar veins. These merge into the renal vein, which exits the kidney through the hilum draining into the inferior vena cava.
Appendix B

Renal Function: Secretion and Re-absorption

During the process of urine formation, blood filtration is followed by secretion and re-absorption of water and other substances in the renal medulla. An overview of renal filtration was given in Section 1.2.1. Re-absorption and secretion are briefly explained below.

B.0.1 Secretion and Re-absorption

Urine formation is a combined process of indiscriminate filtration at the glomeruli followed by selective re-absorption and secretion in the tubules. Thus, for any solute, the amount excreted in urine is the result of the amount of that solute that is filtrated and/or secreted minus the amount of that solute re-absorbed. The precise mechanism is dependent on nature of the substance. Red blood cells and proteins are needed in the blood stream, and thus are not filtered or secreted. Small molecules of substances such as glucose or vitamins, once filtered, are re-absorbed back into the bloodstream. Others substances are unable to cross the filtration barrier and thus are disposed by secretion into the renal tubules. There are others, such as creatinine, that are filtered, secreted, and re-absorbed.

Blood filtration results in the removal of a large proportion of water from the blood-
stream. It is necessary in the ultrafiltrate as the solvent for urine, but most of the water must be re-absorbed as conservation of water levels within the body is an essential function of the kidney. Normal osmolarity\(^1\) in body fluids is 300 mOsm/L. It is the kidneys' function to attempt to maintain that osmolarity by regulating the concentration and the volume of the urine.

Water re-absorption

From the glomerulus, the ultrafiltrate arrives at the proximal tubule, where there is a high volume of re-absorption of both water and solutes (70-80%), since this section of the tubule is lined with highly permeable cells. However, there is a need to re-absorb most of the water in the remaining filtrate. This is achieved by creating an osmotic gradient within the kidney.

Creation of the osmotic gradient. The osmotic gradient is created by the long loops of Henle in the juxtamedullary nephrons and the vasa recta. In the thick ascending limb of these nephrons sodium is re-absorbed into the interstitial fluid, but this segment of tubule is water impermeable. Thus, highly diluted urine (100 mOsm/L) reaches the distal tubule, whereas the sodium re-absorbed in the medulla creates the gradient, with the cortex isosmotic to plasma (300 mOsm/L) and the medulla very hypeosmotic (1200 mOsm/L), compared to plasma.

The vasa recta and the osmotic gradient. Blood in the vasa recta follows a parallel path to urine in the loop of Henle. It looses water and gains sodium as a result of the osmotic gradient as it enters the deep medulla, becoming hypertonic. As opposed to the walls of the thick ascending limb in the nephron, these blood vessels are not water impermeable. As blood flows back to the cortex it gains water and looses sodium, and blood exiting the medulla has the same solute concentration as blood entering the medulla.

\(^1\)Osmolarity is a measure of the osmotic pressure exerted by a solution across a perfect semi-permeable membrane (one which allows free passage of water and completely prevents movement of solute) compared to pure water. Osmolarity is dependent on the number of particles in a solution. It is measured as osmoles per liter of solvent.
The antidiuretic hormone. The final stages of urine concentration occur within the distal convoluted tubule and in the collecting tubule, where water permeability of the walls of these ducts is dependent on the antidiuretic hormone (ADH). In the absence of ADH, low water permeability in these areas results in hyposmotic urine. In the presence of ADH, increased re-absorption of water from the hyposmotic ultrafiltrate into the increasingly concentrated\(^2\) interstitial fluid results in hyperosmotic urine.

\(^2\)Again, as the result of the osmotic gradient between the renal cortex and the renal medulla.
Appendix C

Contrast-Enhanced Magnetic Resonance

C.1 The MRI signal

MRI is an imaging technique based on the Nuclear Magnetic Resonance (NMR) phenomena, whereby the magnetic moments of protons align themselves within a magnetic field. MRI localises and measures the MR signal arising from the hydrogen (\(^{1}\text{H}\)) component in water molecules. The MR signal is dependent on the density of protons and their response to Radio Frequency (RF) pulses: \textit{NMR relaxation}. There are two relaxation processes: \textit{spin-lattice} or \textit{longitudinal} (\(T_1\)) relaxation restores energy equilibrium after the application of a RF pulse and \textit{spin-spin} or \textit{transversal} relaxation (\(T_2\)) destroys phase coherence by the protons instigated by the RF pulse. The resulting MR signal is a complex non-linear function of proton density (PD), \(T_1\) relaxation, and \(T_2\) relaxation. However, the contributions of each of these tissue properties to the MR signal can be controlled by the applied RF pulse sequence. Thus, contrast between different tissues is different in each possible ‘type’ of MR image: \textit{PD}, \(T_1\), or \(T_2\)-\textit{weighted} images. The choice of image to be acquired depends on the properties of the tissue or function of

---

\(^1\text{Hydrogen is the lightest and most abundant element in the universe.}^{1}\text{H} \text{ is its most common form, a stable isotope with a nucleus consisting of a single proton.}^{2}\text{H} \text{ is the other stable isotope of hydrogen, known as deuterium, with an extra neutron in the nucleus.}\)
Appendix C. Contrast-Enhanced Magnetic Resonance

interest.

A further method of controlling contrast in MR is the introduction of contrast agents that change the relaxation properties of local tissues and fluids. Relaxation and the effects of contrast agents on relaxation are briefly described below.

C.1.1 Relaxation

The MRI process consists of the production of an initial magnetisation of protons for alignment with the main magnetic field of the scanner, \( B_0 \), that is in the longitudinal or \( z \)-direction, followed by RF pulses to produce magnetisation in the transverse or \( x-y \) plane. After the application of a pulse sequence, protons fall back into alignment with the external static field \( B_0 \) via the two relaxation processes.

Longitudinal relaxation

Longitudinal relaxation is the process whereby thermal equilibrium is restored after being altered under an RF pulse. This involves a loss of energy of protons into their surroundings and is thus known as spin-lattice relaxation. It is also known as longitudinal relaxation since it corresponds to a restoration of magnetisation along \( B_0 \), the \( z \)-plane. It is characterised by an exponential recovery with a growth constant \( T_1 \), dependent on magnetic field strength (see Figure C.2).

Transversal relaxation

Protons inside a magnetic field \( B_0 \) precess (see Figure C.1 below) at the Larmor frequency, \( \omega_0 \):

\[
\omega_0 = \gamma B_0
\]  
(C.1)

where \( \gamma \) is a constant known as the gyromagnetic ratio which is nuclei specific.
C.1. The MRI signal

Figure C.1: Nuclei are considered to be small spheres of positive charge (due to the presence of protons) that spin and, thus, possess a magnetic moment. If an external magnetic field, \( B_0 \), is applied, the magnetic moment precesses (i.e. the direction of the magnetic moment itself rotates) at the Larmor frequency.

Protons interact with others in close proximity by mutually altering their respective magnetic fields, causing precession at frequencies slightly different to the Larmor frequency. This results in de-phasing. The net result over time and a large number of protons is that all protons become out of phase with each other and thus there is no net magnetic moment. This produces an exponential decay of the signal on the transversal \( x\)-\( y \) plane with time constant \( T_2 \), known as transversal relaxation (see Figure C.2).

However, the spin-spin relaxation described above is not the only contribution to transversal relaxation: inhomogeneities of the magnetic field (\( \Delta B_0 \)) also cause protons to precess at frequencies other than Larmor's. Further, proton diffusion also contributes to transversal relaxation. The combined effects of these and other factors that contribute to de-phasing of protons result in a faster observed decay with time constant \( T_2^* \).

C.1.2 Paramagnetic Contrast Agents

The role of the contrast agent in MR is to enhance the response of tissues to electromagnetic fields. Most common in MRI are paramagnetic contrast agents, such as the naturally occurring deoxyhaemoglobin (oxygen-free haemoglobin present in veins) or
Appendix C. Contrast-Enhanced Magnetic Resonance

Figure C.2: (a) Longitudinal relaxation: at $t = T_1$ after applying an RF pulse restoration of thermal equilibrium reaches 63%; (b) Transversal relaxation: at $t = T_2$ transversal magnetisation decays by 63%. In both cases, spins are relaxed at $\sim 5 \times T$. However, $T_2$ times are considerably shorter than $T_1$ times in biological samples.

exogenous metallic elements such as gadolinium (Gd), and superparamagnetic iron oxide particles (SPIOs). The effect of these contrast agents is to decrease signal intensities on $T_2^*$-weighted images and increase signal intensities on $T_1$-weighted images.

The effects of paramagnetic contrast agents on relaxation times were considered by Solomon (1955), who defined relaxativity $R$, that describes the effect of paramagnetic contrast agents on the relaxation time of water protons. The spin-lattice relaxation time $T_1$ of water in the presence of the widely used contrast agent Gd, is given by:

$$\frac{1}{T_1} = \frac{1}{T_1^0} + R_1 \times [Gd]$$

where $T_1^0$ is the spin-lattice relaxation time of pure water protons in the absence of Gd and $[Gd]$ represents gadolinium concentration. There is ample research in the literature as to which factors, other than simply $[Gd]$, relaxativity is dependent on: the macromolecule to which the contrast agent is attached, temperature, pH, or macromolecular content of the solution (i.e. different for blood and saline solution), have all been found to affect relaxativity (Stanisz and Henkelman, 2000).
The $T_2^*$-shortening Effect

Gadolinium is used to decrease $T_1$ and therefore obtain a signal intensity increase on $T_1$-weighted images. However, in addition to reducing $T_1$, Gd also reduces $T_2^*$. This results in a relationship between measured signal intensity (SI) and [Gd] that is non-monotonic: at high [Gd], the $T_2^*$-shortening effect dominates and produces a non-linear relation between SI and [Gd], with even a decrease of observable SI for very high [Gd] (Brasch et al., 1984; Carvlin et al., 1989; Fichtner et al., 1994).

C.1.3 Conversion of MR Signal to Gadolinium Concentration.

For quantitative MR analysis, such as in MR renography, it is usually necessary to derive from the measured MR signal intensity, $s$ quantitative values for contrast agent concentration, i.e. [Gd]. The simplest approach would be to consider the normalised increase in MR signal intensity to be proportional to [Gd]. Thus:

$$[\text{Gd}] \approx \frac{s - s_0}{s_0} \quad (C.3)$$

where $s_0$ is the baseline MR signal intensity before contrast enhancement (Lee et al., 2003). However, this is not a robust method as signal intensities considerably vary for different patients or different scanner systems (Huang et al., 2003).

A more suitable approach is to model the relationship between the MR signal and the principal parameter being measured, i.e. $T_1$ relaxation in MR renography studies. If $T_1$ can be calculated, then [Gd] can be found by rearranging Equation C.2.

Calculation of $T_1$

Since Gd is a $T_1$ contrast agent, in $T_1$-weighted images $s$ is primarily a function of the longitudinal (spin-lattice) relaxation time, $T_1$, for a given different tissue, $i$:

$$s_i = f_i(T_1) \quad (C.4)$$
and depends on the imaging sequence. The problem is to determine $T_1$ from $s$, that is $f^{-1}$. If $f$ is monotonic, then inversion of Equation C.4 will provide an answer. For some simple sequences $f_i$ can be described analytically. However, heavily $T_1$-weighting sequences often have highly complex formulae, and inversion processes are likely to be non-trivial. In such cases, and for sequences with unknown analytical formulae, $f_i$ has to be estimated empirically (e.g. imaging a phantom consisting of substances with various known $T_1$) (Rusinek et al., 2001).

Calculation of [Gd]

Once $T_1$ is known, then [Gd] can be found by rearranging Equation C.2. This is a simple step but, however, it has been noted that the values of $R$ differ for different solutions or tissues (Stanisz and Henkelman, 2000). This may or may not make a significant difference to [Gd] estimation, but in principle the conversion of $T_1$ to [Gd], is even more challenging than that of $s$ to $T_1$. 
Appendix D

Generation of Synthetic Noise

Synthetic noise is often generated by using a sequence of values from a random source, such as, for example, thermal noise. In scientific experiments, it is often used to be able to replicate results. However, if truly random numbers are used they can not, by definition, be identically re-generated again each time the experiment is run (MacKeeown and Newman, 1987). That is, unless the random numbers used in the first run are stored, which is an impractical solution in most cases. An alternative is to use Pseudo-Random Number Generators (PRNG). These produce a sequence of, hopefully, uncorrelated (random) numbers, which each succeeding random number being generated from the previous one. Thus, knowledge of the initial seed can be used to re-generate the sequence each time.

There are two important parameters in designing/choosing a pseudo-random number generator. Firstly, these sequences generally have a finite period since they repeat themselves after a certain number of realisations. Therefore, in choosing an appropriate random number generator, it is essential that the period be large enough for that particular simulation. Secondly, it is also desirable for correlations between the numbers in the sequence to be as small as possible. Some commonly used pseudo random number generators are briefly described below. First, PRNGs that produce uniformly distributed sequences are presented, followed by a brief review of techniques that are commonly used to transform these uniform distributions into other distributions, such as the Gaussian distribution.
D.1 Uniform Random Numbers

The basic and most useful random number generators are those that produce a sequence of random real numbers \((U_n)\) uniformly distributed in the interval \([0,1]\). That is:

\[
p(n) = \begin{cases} 
1 & 0 \leq n \leq 1 \\
0 & \text{otherwise}
\end{cases}
\] (D.1)

Because of the difficulties in accurately representing real numbers in computers, most methods usually produce a sequence of integers \((X_n)\) between zero and some integer \(m\). In such cases, \(U_n\) can then simply be found by:

\[U_n = \frac{X_n}{m}\] (D.2)

There are many methods for generating this type of uniform random number sequences. For further details on generating random uniform distributions see Knuth (1997).

Linear Congruent Method

Perhaps the most popular type of generator is the Linear Congruent Generator (LCG) first introduced by Lehmer (1951). Using this method, the sequence \((X_n)\) is obtained with the following formula:

\[X_n = (a \times X_{n-1} + c) \mod m \quad \text{for} \quad n \geq 0\] (D.3)

The modulus, \(m\), implicitly defines the largest possible period \((m + 1)\). To speed calculations, \(m\) is normally taken as the word size of the computer (i.e. \(2^{32}\) or \(2^{64}\)). \(X_0\) is the starting value, or seed. The parameters \(a\) and \(c\) are chosen in order to minimise correlation and maximise the period. The case where \(c = 0\), is known as the multiplicative congruential method. The case where \(c \neq 0\), which can produce longer periods, and is known as mixed congruential method.

A good choice of parameters for this formula is extremely important as a badly thought set can quickly lead to correlated short sequences. These type of generators are widely used (the ‘rand()’ function in the ‘ANSI Standard C’ library is a LCG), but they are hampered by small periods, and have difficulties passing stringent randomness tests.
Lagged Fibonacci Generators

Another commonly used type of generator are the Lagged Fibonacci Generators (LFG), using a recursive formula such:

\[ X_n = (X_{n-r} \oplus X_{n-s}) \mod m \]  

where \( s \) and \( r \) are integers for which \( s < r \) and \( \oplus \) denotes binary arithmetic operation, such as addition, subtraction, or multiplication modulo \( m \). The period for these type of generators is \( m^r - 1 \).

Feedback Shift Register Generators

Another important family is that of Feedback Shift Register Generators (FSRG), first proposed by Tausworthe (1965). They can be considered a special case of a LF generator that work on binary registers using the XOR operation and can produce sequences with very large periods. From this group, the Generalised Feedback Shift Register generators (GFSRG) (Lewis and Payne, 1973), are of the form:

\[ x_n = (x_{n-p} \oplus x_{n-q}) \]  

where \( x \) is a binary vector of length \( w \). The maximum period for this type of generator, \( 2^p - 1 \), can be achieved when the trinomial \( x^p + x^q + 1 \) is divisible by \( x^r - 1 \) and \( r \) is a Mersenne prime\(^1\). The Mersenne Twister, proposed by Matsumoto and Nishimura (1998) is a GFSR that includes a matrix used to 'twist' the register and thus improving on the randomness of the generated sequence. The algorithm uses the following formula:

\[ x_{k+n} = x_{k+m} \oplus (x_k^r | x - k + 1^r)A \quad (k = 0, 1, ...) \]  

where \( x \) is again a binary vector of length \( w \), \( x^u \) represents the 'upper' \( w - r \) bits of \( x \) and \( x^f \) represents the 'lower' \( r \) bits of \( x^2 \), and \( | \) denotes concatenation. \( A \) is the \( w \times w \) twisting matrix. The Mersenne twister algorithm published by Matsumoto and Nishimura (1998) has a period of \( 2^{19937} - 1 \), and it is now increasingly becoming the

\(^1\)A Mersenne prime is a prime number \( M(n) = 2^n - 1 \) for which \( n \) is also prime

\(^2\)That is \( x^u = (x_{w-1}, ..., x_r) \) and \( x^u = (x_{w-1}, ..., x_0) \)
'random number generator of choice' for statistical simulations. Its popularity, is based on the large period it can achieve, and the high performance it has achieved when extensively tested.

D.2 Generating Other Probability Distributions

Uniform random number sequences can be used as building blocks for generating new sequences obeying other probability distributions. The main methods for doing this are the Inversion (Direct), Rejection, and Mixed methods, which are presented below. Generation of the widely used Normal distribution is also discussed.

D.2.1 Direct Sampling

The method of direct sampling relies on the definition of a Cumulative Distribution Function (CDF) to map the output of a random number generator as defined in the previous section onto the CDF. The CDF is then inverted to obtain a random sequence with a prescribed Probability Density Function (PDF).

First, a PDF is defined over the interval \([a, b]\) as a non-negative function \(f(x)\) whose integral is equal to 1. That is:

\[
f(x) \geq 0
\]

\[
\int_{a}^{b} f(x) \, dx = 1
\]

Thus, the probability of \(x\) taking on a value in the interval \([x, x + dx]\) is:

\[
F(x) = \int_{x}^{x+dx} f(x') \, dx'
\]

Consider now a CDF, defined as:

\[
F(x) = P(X \leq x),
\]

so that the probability of \(x\) taking on a value less than or equal to \(x\) is:

\[
F(x) = \int_{a}^{x} f(x') \, dx'
\]
By definition, every CDF, $F(x)$, increases monotonically with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1$. This can be used to map a uniform sequence of random numbers in the interval $[0,1]$, $(X_n)$, onto $F(x)$. Thus:

$$X = F(x) \quad (D.12)$$

If the CDF is then inverted:

$$x = F^{-1}(X) \quad (D.13)$$

the result is a sequence of random numbers $(x_n)$ in the interval $[a,b]$ with a probability density function $f(x) = \frac{d}{dx} F(x)$.

The derivative of a CDF is the PDF:

$$\frac{d}{dx} F(x) = f(x). \quad (D.14)$$

---

**Figure D.1: Typical PDF and Corresponding CDF**

This can be visualised as in Figure D.1, by considering two equal intervals $dx_1$ and $dx_2$, with $dx_1$ in a low probability region of $f(x)$, and $dx_2$ near its maximum. These
would correspond to two different intervals in the ordinate of $F(x)$, namely $X_1$ and $X_2$, with their ratio equal to:

$$\frac{dX_1}{dX_2} = \frac{\frac{d}{dx} F(x)|_{x=x_1}}{\frac{d}{dx} F(x)|_{x=x_2}} = \frac{f(x_1)}{f(x_2)}$$  \hspace{1cm} (D.15)

This simple and elegant method only works if $f(x)$ can be integrated (it needs to, by definition, to be a PDF) and $F(x)$ is invertible (either analytically or numerically). This is not often the case and, in practice, although invertible, $F^{-1}$ may be exceedingly difficult to calculate. The Rejection method is then a commonly used alternative.

### D.2.2 Rejection Method

Rejection sampling techniques are used to generate values from an arbitrary PDF, $f(x)$, from an an envelope distribution $e(x)$. Rejection sampling requires $f(x)$ to be bounded in the range $[a, b]$ and its maximum $M$ to be known. Initially, two random numbers $R_1$ and $R_2$ are generated from $g(x)$ (e.g. a uniform distribution). Then calculate $x^* = R_1 \times (b - a) + a$ and $f(x^*)$. If $R_2 \times M \leq f(x^*)$ then $x^*$ is accepted, otherwise it is rejected and a new pair $R_1$ and $R_2$ are chosen.

The acceptance probability is the area under $f(x)$ divided by that of $e(x)$. Clearly, if the area under $f(x)$ is small (e.g. a particularly 'spiky' PDF), this method can prove inefficient.

### D.2.3 Mixed Method

Another method of generating random numbers with a specific PDF is the *mixed method*, as its name implies, it is a combination of the above methods. It consists of factorising the PDF into two (or more) separate functions such:

$$F(x) = g(x) \times h(x)$$  \hspace{1cm} (D.16)

where $g(x)$ is an invertible function which can be sampled using the Direct method, and $h(x)$ is solved using the rejection method. These approaches attempt to maximise efficiency by modelling the 'spiked' regions of $F(x)$ (i.e. those where the rejection method
is inefficient) with an invertible function and the ‘flat’ regions containing the mathematical complexity with a rejection-solved function (i.e. where it would be impossible or to taxing too solve by inversion).

D.2.4 The Box-Muller Transform

The Box-Muller transform is a method of generating pairs of independent normally distributed random numbers. A one-dimensional normal distribution can be defined as:

\[
f(r) = \frac{1}{\sqrt{2\pi}} e^{-\frac{r^2}{2}}
\]  

(D.17)

And its corresponding CDF is thus:

\[
F(r) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{r} e^{-\frac{t^2}{2}} dt
\]  

(D.18)

This CDF is not integrable. However, \( F(r)^2 \) is integrable, through a change of variables into polar coordinates. The Box-Muller transform (Box and Muller, 1958) uses this to transform a pair of uniformly distributed random numbers into a pair of independent normally distributed random numbers. Let \( U_0 \) and \( U_1 \) be independent uniform random numbers distributed in \((0,1]\). Then \( X_0 \) and \( X_1 \) are two normally distributed random numbers such that:

\[
X_0 = \cos(2\pi U_0) \times \sqrt{-2 \ln U_1}
\]  

(D.19)

\[
X_1 = \sin(2\pi U_0) \times \sqrt{-2 \ln U_1}.
\]  

(D.20)

However, this form of the Box-Muller transform can be unstable when \( U_1 \to 0 \), so the polar form of the Box-Muller transform is normally used instead. This form is not only more stable, but faster as there is no need to calculate trigonometric functions.

The polar form of the Box-Muller transform takes two independent uniform random numbers distributed in \([-1,1]\) (which can be simply obtained by applying the following transform random numbers distributed in \([0,1]\): \( Y_u = 2U_u - 1 \). Then set \( R = Y_1^2 + Y_2^2 \). If \( R \geq 1 \), \( R \) is discarded and a new pair of uniform random numbers is calculated.
Otherwise, $X_0$ and $X_1$ are two normally distributed random numbers such that:

$$X_0 = U_0 \times \sqrt{-\frac{2 \ln R}{R}}$$  \hspace{1cm} (D.21)

$$X_1 = U_1 \times \sqrt{-\frac{2 \ln R}{R}}$$  \hspace{1cm} (D.22)

The polar form is thus a type of rejection sampling. Due to the rejection, the calculation of the uniform random numbers is performed, on average, 1.2732 (4/$\pi$) times per output.
Bibliography


A. Blake and M. Isard. *Active Contours: The Application of Techniques from Graph*


J. M. Links, L. S. Beach, B. Subramaniam, M. A. Rubin, J. G. Hennessey, and A. L.


Université de Fribourg. Département de Médecine, Switzerland, last viewed on 8 November, 2005. URL http://www.unifr.ch/histologie/elearningfree/allemand/rein/nier04.html.


