Robust Thin Layer
Coal Thickness Estimation
Using Ground Penetrating Radar

Andrew Darren Strange
B.Eng. (Electrical and Computer Engineering), Hons 1

Speech, Audio, Image and Video Technology Laboratories
School of Engineering Systems
Queensland University of Technology

Submitted as partial requirement for the degree of Doctor of Philosophy
at the Queensland University of Technology
March 2007
Keywords

Abstract

One of the most significant goals in coal mining technology research is the automation of underground coal mining machinery. A current challenge with automating underground coal mining machinery is measuring and maintaining a coal mining horizon. The coal mining horizon is the horizontal path the machinery follows through the undulating coal seam during the mining operation. A typical mining practice is to leave a thin remnant of coal unmined in order to maintain geological stability of the cutting face. If the remnant layer is too thick, resources are wasted as the unmined coal is permanently unrecoverable. If the remnant layer is too thin, the product is diluted by mining into the overburden and there is an increased risk of premature roof fall which increases danger.

The main challenge therefore is to develop a robust sensing method to estimate the thickness of thin remnant coal layers. This dissertation addresses this challenge by presenting a pattern recognition methodology to estimate thin remnant coal layer thickness using ground penetrating radar (GPR). The approach is based upon a novel feature vector, derived from the bispectrum, that is used to characterise the early-time segment of 1D GPR data.

The early-time segment is dominated by clutter inherent in GPR systems such as antenna crosstalk, ringdown and ground-bounce. It is common practice to either time-gate the signal, disregard the clutter by rendering the early-time segment unusable, or configure the GPR equipment to minimise the clutter effects which in turn reduces probing range. Disregarding the early-time signal essentially imposes a lower thickness limit on traditional GPR layer thickness estimators.
The challenges of estimating thin layer thickness is primarily due to these inherent clutter components. Traditional processing strategies attempt to minimise the clutter using pre-processing techniques such as the subtraction of a calibration signal. The proposed method, however, treats the clutter as a deterministic but unknown signal with additive noise. Hence the proposed approach utilises the energy from the clutter and monitors change in media from subtle changes in the signal shape.

Two complementary processing methods important to horizon sensing have been also proposed. These methods, near-surface interface detection and antenna height estimation, may be used as pre-validation tools to increase the robustness of the thickness estimation technique.

The proposed methods have been tested with synthetic data and validated with real data obtained using a low power 1.4 GHz GPR system and a testbed with known conditions. With the given test system, it is shown that the proposed thin layer thickness estimator and near-surface interface detector outperform the traditional matched filter based processing methods for layers less than 5 cm in thickness. It is also shown that the proposed antenna height estimator outperforms the traditional height estimator for heights less than 7 cm.

These new methods provide a means for reliably extending layer thickness estimation to the thin layer case where traditional approaches are known to fail.
# Contents

Keywords i

Abstract iii

List of Figures ix

List of Tables xvi

List of Abbreviations xix

Publications xxi

Authorship xxiii

Acknowledgements xxv

1 Introduction 1
  1.1 The Challenge .............................................. 1
  1.2 Contributions ............................................... 3
  1.3 Scope and Overview ....................................... 3

2 State Of The Art 5
  2.1 Introduction .................................................. 5
  2.2 Sensing The Coal-Rock Interface ............................ 5
    2.2.1 Reactive Methods ...................................... 7
    2.2.2 Predictive Methods .................................... 10
  2.3 Ground Penetrating Radar ................................ 13
    2.3.1 Data Acquisition and Visualisation ................... 14
2.3.2 Layer Thickness Estimation: State-Of-The-Art .................................. 17
2.3.3 Coal Layer Thickness Estimation Using GPR ................................. 18
2.4 Summary ......................................................................................... 22

3 Experimental Data Collection ......................................................... 25
3.1 Introduction .................................................................................. 25
3.2 Experimental Testbed ................................................................. 26
  3.2.1 Design .................................................................................... 26
  3.2.2 Ground Truth for Validation .................................................. 29
  3.2.3 Measurement of Electromagnetic Properties ........................... 31
3.3 GPR Hardware .............................................................................. 31
  3.3.1 Description ............................................................................ 31
3.4 Synthetic Data Generation ............................................................. 35
  3.4.1 The Finite-Difference Time-Domain Method ............................ 35
  3.4.2 Modified FDTD: Real Source Signal ................................. 36
3.5 Summary ......................................................................................... 38

4 Early-Time Feature Vector for Near-Surface Processing .................. 41
4.1 Introduction .................................................................................. 41
4.2 Pattern Recognition ..................................................................... 42
  4.2.1 Pattern Recognition for GPR .................................................. 43
4.3 Higher Order Spectra ................................................................. 45
  4.3.1 Moment and Cumulant Statistics ......................................... 46
  4.3.2 Moment and Cumulant Spectra ............................................ 46
  4.3.3 The Bispectrum ................................................................. 47
  4.3.4 Bispectrum Estimation ....................................................... 49
  4.3.5 The Integrated Bispectrum .................................................. 50
4.4 Proposed Early-Time Feature Vector ....................................... 52
  4.4.1 Segmentation ...................................................................... 52
  4.4.2 Feature Extraction ............................................................ 53
  4.4.3 Class Separation ............................................................... 56
  4.4.4 Noise Reduction ............................................................... 72
4.5 Summary ......................................................................................... 72
7.3.1 Segmentation and Feature Extraction .................................. 130
7.3.2 Antenna Frequency Invariance ........................................ 131
7.3.3 Experiment ............................................................... 132
7.4 Results and Discussion ..................................................... 137
7.5 Summary ................................................................. 144

8 Summary and Future Research .............................................. 145
  8.1 Directions for Future Research ........................................ 147

A Testbed Measurements ....................................................... 149

B Noise Analysis of GPR System ............................................ 153

C Finite-Difference Time-Domain Update Equations ...................... 159
  C.1 Introduction ............................................................. 159
  C.2 2D FDTD Update Equations ........................................ 159
  C.3 Dispersion Model Update Equations ................................ 160
  C.4 Courant Stability Criterion ........................................... 161
  C.5 Absorbing Boundary Conditions ..................................... 161

Bibliography ................................................................. 163
## List of Figures

2.1 Longwall shearer and the two in-seam horizon guidance scenarios. .................. 6
2.2 Sensor positions for the vibration-based CID methods for a continuous coal mining machine. Reprinted, with permission, from [65]. ................................................................. 8
2.3 Continuous miner with an infrared sensor. ©1995 Colorado School of Mines. Reprinted, with permission, from [67]. ................................. 10
2.4 Block diagram of a GPR system. ...................................................... 14
2.5 Figure shows the forms of GPR data visualisation for (a) one, (b) two and (c) three dimensions. .................................................. 15
2.6 C scans of 3D GPR data of buried objects at various depths. .... 16
2.7 Raw scan of FMCW GPR system in a coal mine. ©1978 IEEE. Reprinted, with permission, from [38]. ................................. 20
3.1 Photo of the GPR testbed designed and constructed for ground truth evaluation. ................................................................. 27
3.2 Plan view of the GPR testbed. The testbed has 14 stepped regions and a linear ramp. The layers consist of coal, shale and clay of varying thickness. The bottom layer in all regions is clay. There is no shale in regions 5–10 as these are for investigating coal-clay interfaces. .................................................. 27
3.3 Side profile of the layers for all regions of the GPR testbed. The values shown correspond to the average thicknesses for a given region. 28
3.4 Plan view of the GPR testbed detailing the ground truth measurement points. Region 1 is shown in the top left along with the adjacent regions. ................................. 29
3.5 Plots showing side profile of the testbed area measured at 50 mm spacing with (a) true surface contour and (b) profile referenced to a flat surface. .......................................................... 30
3.6 Measured electrical parameters of coal, shale and clay from testbed from 100 MHz–3 GHz......................................................... 32
3.7 Photos of (a) the GPR electronics and (b) the GPR system with laptop computer used for data acquisition. ......................... 33
3.8 Typical raw A-scan from GPR system ...................... 34
3.9 Real and synthetic GPR traces for ground and air-coupled antennas. The ground-coupled antennas are in contact with a coal surface. 39

4.1 A typical pattern recognition system consists of five main components. The bold components, segmentation and feature extraction, are addressed in this chapter. ................................. 42
4.2 Symmetry properties of (a) the third order moments and (b) the bispectrum. ................................................................. 49
4.3 The bispectrum is integrated along lines with slope $a$ chosen to be spread evenly between 0 and 1. When four parameters are computed, the values for $a$ are 0.25, 0.5, 0.75 and 1. This triangular region is sector 1 in Figure 4.2(b). ........................................... 51
4.4 Window functions used in the feature extraction phase. ........ 53
4.5 Full trace overlayed by segmented signal component. ........ 54
4.6 The magnitude spectrum of a raw GPR signal acquired from the test GPR system....................................................... 55
4.7 FDTD representation of synthetic coal seam floor. ......... 57
4.8 FDTD scan using surface-coupled antennas with top layer only. ... 58
4.9 Figure shows plots of synthetic data with lossless media for (a) changing top layer thickness with fixed permittivity and (b) changing permittivity of bottom layer with top layer thickness of 5 cm. ... 59
4.10 Figure shows plots of synthetic data with conductive bottom layer for (a) changing top layer thickness with fixed permittivity and (b) changing bottom layer conductivity with top layer thickness of 5 cm. 59
4.11 Feature values for rectangular window with length 160 samples showing paths of constant (a) top layer thickness and (b) bottom layer relative permittivity. The data is the same for both plots. 63

4.12 Feature values for Hamming window with length 160 samples with and without sum-of-squares normalisation. 64

4.13 Feature values for rectangular window with 160 samples indicating effect due to varying $a$ parameter. 65

4.14 Intensity representation of Euclidean distances between a given point and the top layer thickness of 10cm point as the window length increases. Features computed with a rectangular window for varying bottom layer relative permittivity. 66

4.15 Feature values for rectangular window with length 160 samples showing paths of constant (a) top layer thickness and (b) bottom layer conductivity. The data is the same for both plots. 68

4.16 Feature values for Hamming window with length 160 samples with and without sum-of-squares normalisation. 69

4.17 Feature values for rectangular window with 160 samples indicating effect due to varying $a$ parameter. 70

4.18 Intensity representation of Euclidean distances between a given point and the top layer thickness of 10cm point as the window length increases. Features are computed with a rectangular window for varying bottom layer conductivity. 71

5.1 Graphical representation of the GPR signal model for layered media with (a) air-coupled and (b) ground-coupled antennas. 77

5.2 Image shows the output of a correlation matched filter as a function of time. Echoes A, B and C are from targets whereas D is due to clutter. 79

5.3 Geometry of layer thickness estimation signal model for ground-coupled antennas. 80

5.4 Geometry of layer thickness estimation signal model for air-coupled antennas. 82

5.5 Raw and background removed B-scan of point reflector. 83
5.6 Raw and background removed B-scan of plane reflector. ........... 83
5.7 Image showing the experimental setup for the coal-shale interface
evaluation. ........................................................................ 85
5.8 Plot showing variation of feature vector for synthetic coal-shale
interface over increasing coal layer thickness (as labelled). ......... 87
5.9 Plot showing feature vector clusters of noisy synthetic data ac-
quired from the coal-shale media. ........................................ 88
5.10 Plot showing true layer thickness vs estimated layer thickness for
synthetic coal-shale interface over increasing coal layer thickness.
The errorbars represent the 99% confidence interval. ............... 89
5.11 Plot showing absolute value of bias of estimator for top layer thick-
nesses of 2 cm and 4 cm. ..................................................... 90
5.12 Plot showing variance of estimator for top layer thicknesses of 2 cm
and 4 cm. ........................................................................ 90
5.13 Plot showing MSE of estimator for top layer thicknesses of 2 cm
and 4 cm. ........................................................................ 91
5.14 Plot showing feature vectors for real data with coal-shale interface
used to train classifier. The labels denote the measured top layer
thickness. ........................................................................ 93
5.15 Plots showing measured vs estimated layer thickness for real coal-
shale interface with increasing coal layer thickness for the proposed
and matched filter based estimators. The error bars represent the
99% confidence interval. ..................................................... 96
5.16 Plots showing the bias for the proposed thin layer thickness es-
imator and the matched filter approach as the SNR is increased
through averaging. ............................................................ 97
5.17 Plots showing the variance for the proposed thin layer thickness
estimator and the matched filter approach as the SNR is increased
through averaging. ............................................................ 98
5.18 Plots showing the MSE for the proposed thin layer thickness es-
timator and the matched filter approach as the SNR is increased
through averaging. ............................................................ 99
5.19 Plot showing the MSE versus true layer thickness for the proposed thin layer thickness estimator and the matched filter approach for no averaging. .................................................. 100
5.20 Image showing the experimental setup for the coal-clay interface evaluation. .................................................. 101
5.21 Plot showing variation of feature vector for synthetic coal-clay interface over increasing coal layer thickness (as labelled). ....... 102
5.22 Feature vectors overlayed for synthetic coal-shale and coal-clay interfaces with various layer thickness. ............................ 103
5.23 Plot showing true layer thickness vs estimated layer thickness for synthetic coal-clay interface over increasing coal layer thickness. The error bars represent the 99% confidence interval. .......... 104
5.24 Plot showing feature vectors generated from real data with coal-clay interface. These values were used to train the classifier. ... 105
5.25 Plots showing measured vs estimated layer thickness for real coal-clay interface with increasing coal layer thickness for the proposed and matched filter based estimators. The error bars represent the 99% confidence interval. ................................. 107
5.26 Plots showing the bias for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging. .................................................. 108
5.27 Plots showing the variance for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging. .................................................. 109
5.28 Plots showing the MSE for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging. .................................................. 110
5.29 Plot showing the MSE versus true layer thickness for the proposed thin layer thickness estimator and the matched filter approach for no averaging. .................................................. 111

6.1 The neural network architecture used for classification. .......... 120
6.2 The bispectral magnitude and phase feature values for varying coal layer thicknesses with (a) coal-clay, (b) coal-shale and shale-clay interfaces. The clusters for no coal, coal thickness less than 5 cm and greater than 5 cm have been circled.

6.3 Detection error tradeoff (DET) curves for the bispectral features, power spectrum and the matched filter for the coal-shale interface.

7.1 Physical model for estimating the height of air-coupled antennas.

7.2 Plot showing feature values for a 500 MHz antenna as the antenna height increases up to one wavelength.

7.3 Plot showing feature values for a 500 MHz, 1 GHz and 1.5 GHz antennas as the antenna heights increase up to one wavelength.

7.4 Diagram of the timber gantry constructed to mount the GPR for the air-coupled measurements.

7.5 Photo of custom built gantry and testbed for air-coupled GPR experiments.

7.6 Photo of GPR system mounted in the air-coupled configuration above the testbed.

7.7 Cubic B spline of measured and interpolated points for classifier.

7.8 True and estimated antenna heights using both matched filter and proposed antenna height estimator.

7.9 True and estimated antenna heights using (a) proposed and (b) matched filter based antenna height estimators. The error bars represent the 99% confidence interval.

7.10 Plots show the bias for the bispectral feature and matched filter based antenna height estimators as the SNR is increased through averaging. The bias is shown for antenna heights of (a) 6 cm and (b) 10 cm.

7.11 Plots show the variance for the bispectral feature and matched filter based antenna height estimators as the SNR is increased through averaging. The variance is shown for antenna heights of (a) 6 cm and (b) 10 cm.
7.12 Plots show the MSE for the bispectral feature and matched filter based antenna height estimators as the SNR is increased through averaging. The MSE is shown for antenna heights of (a) 6 cm and (b) 10 cm.

B.1 Figure shows the residual noise after mean has been subtracted from a free-space scan.

B.2 Figure shows the variance of the additive noise in the raw GPR data with the antennas in (a) free-space, and (b) on the coal surface.

B.3 Skewness, kurtosis and test statistic of the additive noise in the raw GPR data with the antennas in (a) free-space, and (b) on the coal surface.
List of Tables

2.1 Summary of coal-rock interface detection sensors. The desired criteria shown are predictive, real-time, automatic, capable of sensing multiple layers and thickness estimation range covering 0 to 50 cm. Legend: P = predictive, RT = real-time, AU = automatic, L = number of layers, t = thickness. 23

3.1 Range of measured electrical parameters of the testbed layers. The average values over the GPR antenna bandwidth are shown in brackets. 31

3.2 Debye relaxation parameters chosen for FDTD simulations. 36

4.1 Relative permittivity and conductivity values for the synthetic data. The first two columns contain the relative permittivity and conductivity of the top (coal) layer for both experiments. The middle two columns contain the values for the change in permittivity experiment. The last two columns contain the values for the change in conductivity experiment. 57

4.2 Bias and variance statistics of the magnitude and phase features for ensemble averaging and feature averaging. Features were computed from a synthetic GPR signal with additive noise measured from the test GPR system. 73

6.1 GPR testbed regions in feature clusters. 122

6.2 Detection and False Alarm Rates 123

A.1 Measured coal layer thickness for the coal-clay interface evaluation of the thin layer thickness estimator. 150
A.2 Measured coal and shale layer thickness for the coal-shale interface evaluation of the thin layer thickness estimator. . . . . . . . . . . 151
A.3 Measured thickness for coal and shale layers, and clay interface depth. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 152
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CID</td>
<td>Coal-Rock Interface Detection</td>
</tr>
<tr>
<td>FDTD</td>
<td>Finite-Difference Time-Domain</td>
</tr>
<tr>
<td>FMCW</td>
<td>Frequency-Modulated Continuous Wave</td>
</tr>
<tr>
<td>GPR</td>
<td>Ground Penetrating Radar</td>
</tr>
<tr>
<td>HOS</td>
<td>Higher Order Spectra</td>
</tr>
<tr>
<td>MSE</td>
<td>Mean Squared Error</td>
</tr>
<tr>
<td>NGR</td>
<td>Natural Gamma Radiation</td>
</tr>
<tr>
<td>PRF</td>
<td>Pulse Repetition Frequency</td>
</tr>
<tr>
<td>RMPA</td>
<td>Resonant Microstrip Patch Antenna</td>
</tr>
<tr>
<td>UXO</td>
<td>Unexploded Ordnance</td>
</tr>
</tbody>
</table>
Publications


Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: _______________________

Date: _______________________

xxiii
Acknowledgements

I would firstly like to thank my principal supervisor Assoc. Prof. Vinod Chandran for his support, guidance and endless patience throughout my candidature. I couldn’t ask for a more supportive supervisor. Thanks also goes to Prof. Sridha Sridharan for the financial support provided throughout my time at QUT. I would also like to acknowledge the timely words of wisdom by my associate supervisor, Dr. Jonathon Ralston. Your input during the final stages was invaluable.

Thanks goes to my employer, Con Caris, for providing latitude, encouragement and time allowances to enable me to complete this dissertation while still working full-time. I also acknowledge the support of the CSIRO Mining ICT and Automation group for the use of their GPR equipment, facilities and funding of the testbed and associated materials. In particular, thanks goes to Andrew Castleden for his efforts in constructing the testbed and gantry.

Thanks also goes to the members of the QUT speech lab for the enjoyable times throughout my time at QUT. In particular, thanks goes to Dr. Michael Mason for introducing me to the world of neural networks and pattern recognition.

I would also like to thank my parents and family for continually being interested in my progress, Graham and Chris for the initial motivation to consider tertiary study, and my friends at North Brisbane Church for continually asking am I finished yet?

Most importantly, I wish to sincerely thank my wife, Lisa, for her unconditional love, emotional support and encouragement throughout the last 9 years to finally realise this goal. And to my two fantastic boys, Hunter and Campbell, we’ll now be able to spend lots of time together. My final word of thanks goes to my Lord and Saviour, Jesus Christ.
Chapter 1

Introduction

1.1 The Challenge

One of the most significant goals in coal mining technology research is the automation of underground coal mining machinery. Amongst the many important automation tasks, one largely unsolved problem is measuring and maintaining a coal mining horizon [85] (this is called horizon control in the coal mining industry). The coal mining horizon is the horizontal path the machinery follows through the undulating coal seam during the mining operation. A typical mining practice is to leave a thin remnant of coal unmined in order to maintain geological stability of the cutting face. If the remnant layer is too thick, resources are wasted as the unmined coal is permanently unrecoverable. If the remnant layer is too thin, the product is diluted by mining into the overburden and there is an increased risk of premature roof fall which increases danger. The overall benefits of a sensing method to realise automatic horizon control are improved coal quality, increased mining productivity and enhanced worker safety [88].

The main challenge therefore is to develop a robust sensing method to estimate the thickness of thin remnant coal layers. For future integration into an automatic control system, three key requirements have been identified, namely (1) a reliable means to determine if the remnant layer is thin; (2) a reliable means to estimate the thickness of the thin layer; and (3) a reliable means to estimate the height of the sensor above the surface.
This dissertation presents a pattern recognition methodology to estimate thin remnant coal layer thickness using ground penetrating radar (GPR). The approach is based upon a novel feature vector, derived from the bispectrum, that is used to characterise the early-time segment of 1D GPR data.

The early-time segment is dominated by clutter inherent in GPR systems such as antenna crosstalk, ringdown and ground-bounce. It is common practice to either time-gate this signal, disregard it by rendering the region unusable, or configure the GPR equipment to minimise these effects which in turn reduces probing range. This is an unfortunate practice as it removes the very portion of the signal that is needed for thin layer coal thickness estimation. Disregarding the early-time signal essentially imposes a lower thickness limit on traditional GPR layer thickness estimators.

The challenges of estimating thin layer thickness is primarily due to these clutter components. Traditional processing strategies attempt to minimise the clutter using pre-processing techniques such as the subtraction of a calibration signal. The proposed method, however, treats the clutter as a deterministic but unknown signal with additive noise. Hence the proposed approach utilises the energy from the clutter and monitors change in media from subtle changes in the signal shape.

Two processing methods that complement the thin layer thickness estimator have been proposed. These methods, near-surface interface detection and antenna height estimation, may be used as pre-validation tools to increase the robustness of the thickness estimation technique. The near-surface interface detector determines if the top layer is actually thin. The antenna height estimator determines how close the antenna is to the surface or if the antenna is in contact with the surface.

The proposed methods have been tested with synthetic data and validated with real data obtained using a low power GPR system and a testbed with known conditions.
1.2 Contributions

The significant original contributions of the dissertation are as follows.

1. Development and quantitative analysis of bispectral-based features to characterise the early-time signal of 1D GPR data.

2. Development of a processing method to estimate the thickness of a thin top layer with GPR.

3. Development of a processing method to detect the presence of a near-surface interface with GPR.

4. Development of a processing method to estimate the height of a GPR antenna which includes the height range of less than one third of the electromagnetic wavelength.

Minor contributions include: the design and construction of a testbed for generating data to validate the proposed methods; the measurement of the electromagnetic properties of coal, shale and clay samples; the adaptation of the finite-difference time-domain (FDTD) method to generate synthetic data which incorporates antenna ringdown; and the noise evaluation of a low power GPR system.

1.3 Scope and Overview

Chapter 2 introduces the need for layer thickness estimation in underground coal mining and presents a review of coal-rock interface sensing methods. A review of GPR is also presented which discusses layer thickness estimation and prior research into coal layer thickness estimation with GPR.

Chapter 3 describes the primary source of data used in the experiments. Firstly, the design and construction of a testbed for validating the proposed processing strategies is discussed. This includes the measurement of the electromagnetic parameters of the testbed media. The details of a low power GPR system for real data acquisition are presented along with the FDTD method implemented for synthetic data generation. A modification to the FDTD
method is described that allows the modelling of a real GPR system in the FDTD domain.

Chapter 4 presents the proposed feature vector that characterises the early-time signal of 1D GPR data. This feature vector is central to the thin layer thickness estimator, near-surface interface detector and antenna height estimator proposed in Chapters 5, 6 and 7 respectively. A review of higher order spectra, with special attention given to the bispectrum, is presented along with a review of pattern recognition work applied to GPR.

Chapter 5 introduces a novel thin layer thickness estimation processing method. The technique is a pattern recognition approach utilising the early-time feature vector presented in Chapter 4 and a nearest neighbour classifier. Results are presented from both synthetic and real GPR data.

Chapter 6 describes the first pre-validation processing method – near-surface interface detection. The objective is to determine if the top layer is thin by detecting the presence of an interface within a given distance of the surface. The approach utilises the early-time feature vector and a neural network classifier. The method has been compared with features derived from the power spectrum and a matched filter detector. Results from real GPR data are presented.

Chapter 7 describes the second pre-validation processing method – antenna height estimation for antennas close to the surface. This chapter shows that the proposed antenna height estimator is reliable up to an antenna height of one-third of the electromagnetic wavelength, as compared with the traditional matched filter approach that fails in this region.

Chapter 8 presents a summary of the dissertation and outlines possible paths for further research related to this field.
Chapter 2

State Of The Art

2.1 Introduction

This chapter presents a state-of-the-art review of coal-rock interface detection and layer thickness estimation for underground coal mining applications. A review of ground penetrating radar is also presented. This sensor has shown promise for the interface detection and layer thickness estimation tasks in underground coal mining.

2.2 Sensing The Coal-Rock Interface

One of the current challenges with automating underground coal mining machinery is measuring and maintaining a coal mining horizon [85]. The coal mining horizon is the horizontal path the machinery follows through the undulating coal seam during the mining operation. There are usually two in-seam horizon guidance scenarios depending upon the coal seam geology [94]. The first is to extract all of the coal right to the seam boundary as in Figure 2.1(a). The second is to leave a thin remnant coal layer unmined on the roof, floor or both as in Figure 2.1(b). For the latter case, there is an optimal remnant thickness determined by site conditions and is often chosen when a thin layer of low quality coal due to higher levels of ash and sulphur exists between the actual coal seam and surrounding rock layers [102]. If the remnant layer is too thick, resources are wasted
Chapter 2. State Of The Art

This is the longwall shearer mining right down to the roof and floor overburden.

(a) Coal extracted to the seam boundary

(b) Coal extracted with remnant coal layer on the roof and floor

Figure 2.1: Longwall shearer and the two in-seam horizon guidance scenarios.

as the unmined coal is permanently unrecoverable. If the remnant layer is too thin or the miner cuts into the surrounding strata, the extracted coal is diluted and there is an increased risk of roof fall. The mining challenge therefore is to develop a sensing method to detect the interface between the coal and surrounding layers. This sensing method is called coal-rock interface detection (CID). In addition to detecting the interface, the complementary task of estimating the remnant coal seam thickness provides more information to facilitate intelligent guidance decisions.

Existing interface sensing methods for the underground coal mining industry can be separated into two categories – reactive and predictive. Reactive sensors detect the interface when the machinery penetrates the overburden. Predictive approaches however sense the remnant coal thickness before it is mined and thus allow for optimal mining.

The benefits of predictive sensing and leaving an appropriate remnant coal
layer are: (1) improved coal quality as a result of not extracting coal that has higher levels of impurities such as sulphur and ash which is common in coal close to the overburden; (2) increased mining productivity by extracting the optimal amount of coal and not requiring to slow down extraction when approaching the seam boundary; and (3) enhanced worker safety by removing the operators from the hazardous mining face [65].

This section discusses these interface sensing strategies.

### 2.2.1 Reactive Methods

Reactive sensors are based on detecting changes in the mining operational characteristics when the coal-rock interface is encountered. The reactive sensors are sub-optimal as the mining machinery has already cut into the rock when the interface is detected which damages the machinery and dilutes the coal.

Many different sensing methods have been investigated since the concept of coal-rock interface detection was conceived, many of which can be considered reactive approaches. A detailed review of all of the reactive methods is not justified for inclusion in this dissertation as this approach is fundamentally limited in nature. Work by NASA [71] and Mowrey [64, 66] provide details on other reactive sensing approaches. As such, three types of reactive sensors – cutting drum motor current, vibration [88] and infrared thermography [67] – are described.

**Cutting Drum Motor Current**

The cutting drum motor current sensor has been applied to highwall coal mining operations to detect coal-rock interfaces. The sensor monitors the current driving the rotating cutter drum. When the machinery is mining coal, the current is typically within a certain range of values. As the miner progresses beyond the coal into the overburden, the current required to drive the drum through the harder geology increases. This change in current is a suitable indicator of when the mining operation has departed from the coal seam [66].
Chapter 2. State Of The Art

Figure 2.2: Sensor positions for the vibration-based CID methods for a continuous coal mining machine. Reprinted, with permission, from [65].

Vibration

Three vibration-based CID techniques were investigated by the United States Bureau of Mines (USBM) for continuous coal mining machinery – the in-seam seismic, mining machine vibration and acoustic methods [65]. These vibration methods are based on the principle that each mining operation (cutting coal or cutting rock) has a distinct vibration signature. As described in [65], this is primarily due to differing mechanical properties of the mine geology being cut.

The in-seam method monitors seismic vibration signals acquired from one or more accelerometers mounted on roof bolts, the coal seam, and/or exposed roof or floor strata behind the mining operation [63]. The mining machine method measures mechanical vibration of the cutter head using accelerometers mounted at various points on the machinery. The latter method measures acoustic vibration signals from the mining environment using a sound pressure sensor (microphone) [65]. Figure 2.2 shows an example sensor position configuration for the three vibration-based CID methods.

The advantages of the in-seam method are less machinery noise in the data and an immediate seismic response when the cutter breaks through the coal-rock interface. The disadvantages however are the signal dependency on the distance to the vibration source and the sensors must be moved periodically to keep up with the mining operation [62].
The advantages and disadvantages of the mining machine vibration method are converse to that of the in-seam method. The accelerometers are in a fixed location hence the need to continually move and remount the sensors is avoided. There is however decreased signal-to-noise ratio (SNR) due to higher noise levels from the machinery [62]. Additionally, the sensors must be able to withstand the harsh environment directly at the mining operation.

The acoustic sensors can be mounted to the coal seam walls similar to the in-seam seismic method or on the mining machine. The advantages and disadvantages for this method are determined by the mounting configuration chosen [65]. One other disadvantage is the mining environment is acoustically loud due to the mining operation resulting in increased noise in the acoustic data.

**Infrared Thermography**

Another reactive CID strategy investigated by the USBM is infrared thermography. Similar to the cutting drum motor current and vibration-based techniques, the infrared thermography CID method relies on the contrasting mechanical properties of the coal layer and host strata. When the coal-rock interface is encountered, the temperature of the cutter picks and nearby geology rise due to the increased friction from cutting overburden. During experiments conducted by Mowrey and Ganoe [67], it was deduced that a typical cutter pick temperature range is 25–32°C when cutting coal and 50–60°C when cutting overburden. The sensing method utilises a thermal-imaging infrared camera to generate an image of the cutter drum and surrounding geology temperatures (see Figure 2.3) and a video motion detector for automatically detecting changes in the image. This system was developed for future incorporation into an automated horizon control system.

One advantage of infrared sensors is they can usually see through the dust and water spray inherent in an underground coal mining environment. As a result, the infrared sensor will function without the need for constant cleaning as would be required for visible light cameras [94]. However if the dust cloud or water is too dense, the temperature change cannot be detected and the infrared sensor fails to detect the interface [67].
The infrared sensor selected for testing was the Pyroviewer Model 5400 [67]. This sensor uses a vidicon tube that has pyroelectric material to detect the infrared energy. This sensor was considered relatively rugged as it does not contain any moving parts. So it is suited to the mining environment in this respect. One drawback with the design is that pyroelectric material is sensitive to mechanical vibration hence images generated by this sensor will be susceptible to additive noise when subjected to vibration [67].

The infrared sensor mounted inside an explosion-proof box and video motion detector were tested in a real coal mining operation to detect the coal-rock interface for an automatic control system. The results were not as successful as initial laboratory tests. Mowrey and Ganoe [67] concluded that the poorer results were attributed to inadequate focusing of the sensor and added electrical noise from the mining machine vibration. However the method did show promise as a reactive CID sensor and potential for further investigation.

2.2.2 Predictive Methods

Predictive CID approaches are an improvement over reactive strategies. They estimate the remnant coal thickness before mining has occurred. There are four sensors that fit into the predictive category. These are the natural gamma radiation sensor, optical marker band detector, the electromagnetic wave detection and imaging transceiver (EDIT) and ground penetrating radar. These sensing
methods are discussed in the following sections.

Natural Gamma Radiation

The most common and well established predictive CID method uses a natural gamma radiation (NGR) sensor to estimate the remnant coal thickness by exploiting the radioactivity of host strata found with many coal deposits. NGR sensors have been successfully commercialised for the interface sensing task and can estimate remnant coal thickness over the range 2–50 cm [65].

Many overburden materials such as shale, clay, silt and mud contain very small traces of radioactive elements such as potassium (K-40), uranium and thorium. Gamma particle emissions from shale type geology can typically range from 10–40 picocuries per gram. In contrast coal is usually less than 2 picocuries per gram [65]. As a result of this natural phenomenon, the overburden is considered a NGR source and the coal a NGR absorber. An estimate of the remnant coal layer thickness is obtained from the number of emitted gamma particles counted over a set time duration compared with that of the exposed host strata during sensor calibration.

Even though the application of a NGR sensor for predictive CID sensing is theoretically sound, there are still practical limitations with the method. Firstly, radioactive decay is a random process hence the accuracy of the coal thickness estimate is proportional to the number of counts recorded. To increase the estimate accuracy, either a very large sensor or long data collection times are required [123]. Long counting times of up to 30 seconds can severely impact coal mining operations and reduce production rates [88].

Secondly, the successful application of the NGR sensor is highly site specific. Non-shaley type geology such as sandstone and limestone have lower levels of radioactive elements. The NGR sensor therefore will not work in coal seams with these overburden materials [64].

These practical limitations of the NGR sensor motivate the search for alternative coal thickness estimation techniques.
Chapter 2. State Of The Art

Optical Marker Band Detection

Optical marker band detection is based on the way operators currently perform predictive coal-rock interface detection manually [50]. As coal seams are layered geology, often there exists visual markers in the form of bands in the seam profile. During the mining process, the operators raise or lower the shearer drum to track a specific marker band. The operators can infer the remnant coal seam thickness based on a previously measured distance (from drilling) between the marker band and the coal-rock interface. This process can be automated using image sensors however the problem is the remnant coal thickness is not actually being measured continuously. Rather the thickness estimate is based on a previously measured distance between the marker band and the actual coal-rock interface. Furthermore, the image sensors or their housing units require constant cleaning due to the dusty conditions from the mining operation.

Electromagnetic Wave Detection and Imaging Transceiver (EDIT)

Stolarczyk et al. [104] developed and patented a sensor for estimating the thickness of the first layer of a geologic deposit (e.g., coal, trona and potash). The thickness estimation range is from 6–300 mm.

The sensor operates on the phenomenon that the conductance and resonant frequency of a resonant microstrip patch antenna are sensitive to changes in the first layer thickness of a medium being scanned. The measured conductance and/or resonant frequency are mapped to a set of recorded values obtained during a calibration process. The thickness is then estimated by interpolating between the values measured during calibration to values measured during the test.

This system was installed for trials on a continuous miner in Carlinville, Illinois, USA [105]. However, to the author’s knowledge, the outcome of these trials are yet to be published in the literature.
2.3. *Ground Penetrating Radar*

GPR holds several fundamental advantages over the other predictive sensors. Firstly, the acquisition rate of raw GPR data is very fast as electromagnetic energy propagates through coal at approximately 15 cm/ns. Hence any limitations in data acquisition rates are due to hardware limitations, not physical limitations that cannot eventually be overcome through advancing technology (*i.e.*, compared with the gamma particle emission rate for the NGR sensor which is a comparatively slow process). Secondly, the thickness is estimated by measuring the electromagnetic wave propagation time to the target interface (as opposed to predicting the thickness from the presence of a marker band). Finally, unlike the EDIT sensor, GPR is capable of sensing multiple interfaces. This is important when banding is present in the coal seam or overburden.

One significant disadvantage however with the use of GPR for remnant coal thickness estimation is the fact that layers thinner than the GPR system resolution are detected as a single interface, not as separate upper and lower interfaces. In the case of a thin remnant coal layer, the reflection from the coal-rock interface is masked by clutter components inherent in GPR systems. Hence the coal-rock interface will be undetectable and the miner may continue to penetrate the coal and eventually rock when it should retreat. Overcoming this significant limitation of undetectable thin remnant coal seams is one of the main motivating factors of this dissertation and will be discussed in later chapters. The next section presents an overview of GPR and includes a review of prior research of coal layer thickness estimation using GPR.

**2.3 Ground Penetrating Radar**

GPR is a non-invasive sensing technique used to obtain information about media in the earth’s surface. In impulse GPR systems, a short pulse of electromagnetic energy is transmitted into the ground. A proportion of this energy is reflected back towards the surface at interfaces of media with differing electromagnetic properties. Information such as layer thickness can be extracted from the amplitude and time delay of these reflections. There are many applications that use
GPR for sub-surface imaging such as archaeology, pipe and cable detection, geological profiling, buried landmine detection [31], snow thickness estimation and pavement evaluation [81]. The block diagram of a generic GPR system is shown in Figure 2.4.

![Block diagram of a GPR system.](image)

**Figure 2.4: Block diagram of a GPR system.**

### 2.3.1 Data Acquisition and Visualisation

GPR data can be acquired in either one, two or three spatial dimensions. These datasets are referred to as the A-scan, B-scan and C-scan respectively. 1D GPR data for A-scans is acquired with the GPR antennas stationary in a fixed position. The most common form of A-scan visualisation is a signal amplitude versus time plot as in Figure 2.5(a). Processing and/or knowledge of the electromagnetic properties of the surface can enable the conversion of the time axis into a depth scale.

2D GPR data for B-scans is collected by moving the antennas with a fixed separation offset along a line acquiring data at specific points. This can be regarded as multiple A-scans lined up beside each other. The B-scan is often visualised using an image where each A-scan trace is represented as a column of pixels. The signal amplitude of each trace is mapped to either a greyscale intensity or an
2.3. Ground Penetrating Radar

index into a colourmap. It is common practice to use colourmaps that aid visual data interpretation. These images are also called linescans. Figure 2.5(b) shows a B-scan image.

3D GPR data for C-scans is obtained by repeatedly acquiring B-scans offset by a pre-determined distance. The most common form of visualising C-scans is with horizontal slices through the data as shown in Figure 2.6. With additional processing, 3D GPR data can be displayed using volume visualisation.
Figure 2.6: C scans of 3D GPR data of buried objects at various depths.
2.3.2 Layer Thickness Estimation: State-Of-The-Art

The target of interest for the coal thickness estimation application is the interface. One application for GPR that has a target type as an interface is pavement evaluation. The use of GPR for estimating the thickness of sub-surface pavement layers has been well documented [59, 89, 91, 101] and is a well established technique. Current techniques for this application were quantitatively tested with both asphalt and concrete pavements.

Asphalt pavement thickness can be reliably estimated over the range 50–200 mm with an accuracy of 2.5 mm with site calibration [59]. This method uses a GPR system with bistatic air-coupled horn antennas mounted to the back of a vehicle. These systems can acquire data at a rate such that the vehicle can travel at highway speeds scanning the pavement without the need for lane closure. Information about the pavement can be obtained from the peak amplitude and time delay of reflected echoes, or by the use of the matched filter [89, 101].

In contrast to asphalt pavement which typically has low conductivity, concrete pavement is lossy due to the water content [59]. As a result of the signal attenuation from propagating through a lossy medium, the air-coupled GPR system for asphalt evaluation is unable to estimate concrete thickness. This led to the investigation of ground-coupled antenna systems. In this configuration, the antennas are in direct contact with the surface which allows more energy to propagate into the ground. There are two survey techniques in which ground-coupled antennas are used. The more common for qualitative real-time mapping is when the transmitter and receiver antennas are in the fixed-offset bi-static configuration. Maser [59] reported that the accuracy of thickness estimates obtained with the fixed-offset ground-coupled antennas is not satisfactory for either asphalt or concrete pavement evaluation. This is primarily due to the uncertainty in the wave propagation velocity estimate of the top layer.

The second survey technique used with ground-coupled antennas is the common mid-point. This method showed reasonable results for estimating concrete pavement thickness however the minimum detectable thickness was 75 mm. Unfortunately the common mid-point method cannot be used to acquire data at highway speeds as the antenna separation needs to be progressively increased
between scans while scanning the same point.

A common signal processing tool used to detect target echoes in GPR data is the matched filter. The limiting factor with using the matched filter for detecting reflections from shallow interfaces is the clutter inherent in GPR systems. For the ground-coupled antennas case, the near-surface target reflection is masked by the crosstalk and antenna ringdown. The crosstalk consists of the direct transmitter to receiver signal. The antenna ringdown is caused primarily by the re-radiated fields due to currents reflecting within the antenna and associated structures [5]. One solution to minimise the clutter is to use air-coupled antennas where the antenna ringdown and crosstalk have attenuated before target reflections are received. However in this case, the near-surface target reflection can be masked by the ground-bounce reflection. Furthermore, the use of air-coupled antennas results in a decreased probing range.

A common technique used to remove the clutter is background subtraction. This involves subtracting the mean trace from the data. Background subtraction can give satisfactory results when imaging isolated objects such as buried landmines, however the response from an interface is often filtered out using this process.

Other processing schemes typically suppress clutter for imaging point reflectors or detecting landmines [1, 39, 93, 121, 122], or involve forward and inverse modeling based on physical models and approximate numerical solutions to Maxwell’s equations [55, 56]. Although these methods have shown good results at imaging landmines and characterising the subsurface, they can be difficult to apply to different imaging scenarios and can be computationally intensive. Computationally intensive processing methodologies such as those based on inversion and optimisation are impractical for a real-life mining scenario. The delays encountered during processing will severely limit production.

2.3.3 Coal Layer Thickness Estimation Using GPR

The use of GPR for probing coal has been investigated by various researchers over the past 30 years. GPR was first considered as a sensor for coal mining applications by Cook [27] who measured the complex permittivity of several geological
2.3. Ground Penetrating Radar

Materials including coal at the frequencies 1 MHz, 5 MHz, 25 MHz and 100 MHz. It was discovered that coal is relatively translucent at GPR frequencies [27].

The first reported case where GPR was used to estimate remnant coal layer thickness as a potential sensor for an automated mining machine was in 1974 by Ellerbruch and Adams [37] of the National Bureau of Standards. A frequency-modulated continuous wave (FMCW) GPR system was developed with a frequency range of 1–2 GHz and broadband rectangular aperture horn antennas. The system was tested in the laboratory and three coal mines with reasonable results [38].

The propagation time delay of FMCW GPR systems is estimated from the change in centre frequency between peaks in the received signal spectrum. Figure 2.7 shows a raw scan acquired using the FMCW system in a coal mine with a remnant coal layer beneath a shale ceiling. The main peak is the reflection from the air-coal interface and the peak marked with the dashed line is from the coal-shale interface. One facet not addressed in this research was a method of automatically distinguishing between target interface reflections and those from anomalies such as that labelled “top binder”.

Roe and Wittmann [90] continued the work by Ellerbruch and Belsher by developing another FMCW GPR that operated in the frequency range of 1-4 GHz. The objective of this project was to build on the earlier work and develop an automatic system for coal layer thickness estimation. A testing facility was constructed utilising blocks of material with similar electrical properties of coal to evaluate system accuracy. The continued development of the system was ceased in 1981 as the FMCW GPR was too sensitive to changes in antenna height from the air-coal interface. This sensitivity may have been overcome with system calibration however this would render the system impractical for continuous use. Additionally, further research to overcome these limitations was not warranted as the NGR sensor described in section 2.2.2 was already in operation for the coal thickness estimation task with sufficient accuracy [90].

Daniels [29] investigated the feasibility of remnant coal layer thickness estimation with a 1 GHz impulse GPR and air-coupled antennas. Reflections from a coal-shale interface with coal layer thicknesses of 10 and 20 cm were visible with
this system. The concept of subtracting a calibration signal from the raw data (background subtraction) was proposed to enhance interface detection. It was concluded that it is feasible to use GPR for the coal layer thickness estimation task and as a sensor for an automatic control system, however more sophisticated signal processing techniques were required [29].

Chufo and Johnson [26] of the USBM developed a stepped-frequency GPR system to estimate the dielectric constant and layer thickness of a remnant coal layer. The system was tested in a laboratory and validated in an underground coal mine. The thickness of a 15.2 cm (6 inches) coal seam was estimated as 13.5 cm (5.3 inches).

A unique spatial modulation technique was developed which involved shifting the antenna perpendicular to the coal surface in 32 steps from 10–50 cm. The spatial modulation technique was developed to reduce multipath effects. At each spatial step, the reflection coefficient was measured at 401 frequencies over the range 0.6–1.4 GHz.

A model based on one-dimensional spherical wave scattering matrix theory was proposed. The model calculated the frequency dependant complex reflection coefficient from defined layer thickness, dielectric constant values and antenna offset positions. The stepped-frequency GPR system measured the reflection coefficients over the nominated frequency range and antenna to surface offsets. The measured and modelled reflection coefficients were correlated. The thickness
parameter of the modelled reflection coefficient with the highest correlation to the measured data was the chosen result [26].

Even though this system was able to estimate coal layer thickness, it is unsuitable as a sensor for the underground mining application. Firstly, it is too slow due to the time required to repeatedly move the antenna and acquire the data for a single estimate. Secondly, it is likely that the antenna positioning module would not survive the harsh underground mining environment.

Mowrey et al. [68], also with the USBM, tested a commercial impulse GPR system for monitoring coal rib thickness in highwall mining. The Geophysical Survey Systems Inc. SIR-10 system was chosen with a 500 MHz centre frequency bow-tie antenna module. The rib in highwall mining is the unmined coal between adjacent tunnel entries to support the roof. Therefore the reflections received by the GPR system are from air-coal and coal-air interfaces. The rib thickness was estimated from the time delay between the interface reflections and an assumed propagation velocity [68].

It was concluded that for best results, the coal surface being scanned should be relatively clean and the antennas should be within 5 cm of the surface. The results from this research were promising as the system was found to estimate rib thickness from 0.9 m to more than 3 m.

In 1996, the results of a three year research program sponsored by the Australian coal industry were published by Yelf [127]. One of the objectives of the project was to demonstrate the capability of GPR to monitor changes in the coal-rock interface in real time to assist operators with machine guidance. Band pass filters and a smoothing technique were implemented to remove noise and jitter from the data. The data was presented to mine operators in the form of a seam profile [127] however no automatic thickness estimation processing for automatic system control was investigated.

A patent was filed in 1996 by Stolarczyk and Stolarczyk [103] to protect the commercial implementation of a multi-sensor system for estimating coal seam thickness. The proposed system was based on the EDIT sensor (section 2.2.2) operating in conjunction with a stepped-frequency GPR. This system measured the reflected electromagnetic energy at 64 equally spaced frequency steps over
the frequency range 200–1000 MHz. The antenna module consisted of two circularly but oppositely polarised patch antennas to reduce the direct transmitter to receiver signal.

As mentioned previously in section 2.2.2, the maximum thickness range of the EDIT sensor was 30 cm. The proposed stepped-frequency GPR system was designed to only estimate coal layer thickness above 30 cm [103]. Although this proposed system with two sensors operating simultaneously covers both thickness ranges, to the author’s knowledge, there is no other published work on this concept in the literature.

Ralston and Hainsworth of CSIRO [85, 86, 87, 88] investigated coal seam thickness estimation using a custom built impulse GPR. The GPR system had a depth range of up to 30 cm in coal. A proof-of-concept project was conducted to evaluate whether the electrical parameter contrast in the coal mine strata was sufficient for GPR to be successful at detecting the coal-rock interface. Pre-processing algorithms were proposed for the coal thickness estimation task. These included exponential time-varying gain to compensate for propagation losses, background subtraction to remove the direct transmitter to receiver signal, ensemble averaging to reduce noise and wavelet deconvolution to enhance signal detection [85]. GPR was deemed a promising sensor for the underground mining application however the thickness estimation range via a quantitative study was not evaluated.

2.4 Summary

There are two general in-seam horizon guidance approaches for extracting product from an underground coal mining operation. The first is to mine right to the overburden extracting all of the coal. The alternative and preferred method is to leave a thin layer of coal unmined. The significant advantages with this approach are improved coal quality, increased mining productivity and increased safety for mining personnel.

For a real mining environment, a sensor that is practically viable for sensing coal-rock interfaces should have the following attributes:
2.4. Summary

1. Predictive or non-invasive,
2. Operates in near real-time,
3. Automatically generates defined output without the need for interpretation by trained personnel,
4. Capable of sensing multiple layers, and
5. Covers thickness range from 0–50 cm.

A summary of the state with reference to these criteria for the sensors reviewed in this section is shown in Table 2.1.

<table>
<thead>
<tr>
<th>Method</th>
<th>P</th>
<th>RT</th>
<th>AU</th>
<th>L &gt; 1</th>
<th>0 &lt; t &lt; 50 cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motor current</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>Seismic vibration</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>Machine vibration</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>Acoustic vibration</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>Infrared thermography</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
<tr>
<td>NGR</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Optical marker band</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EDIT</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPR</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Summary of coal-rock interface detection sensors. The desired criteria shown are predictive, real-time, automatic, capable of sensing multiple layers and thickness estimation range covering 0 to 50 cm. Legend: P = predictive, RT = real-time, AU = automatic, L = number of layers, t = thickness.

Of the predictive sensors, the first and most well established is the NGR sensor. Unfortunately, this sensor has processing delays which can reduce production rates and it is not suitable for certain geology. Although the EDIT sensor is capable of estimating uncut coal thickness from 6–300 mm, it is unable to image or detect multiple layers. The optical marker band detection strategy bases its thickness estimate on previously measured values. The coal layer thicknesses are not uniform as assumed by this approach.

GPR has been investigated sporadically for the coal thickness estimation task for 30 years with varying levels of success. The GPR hardware technology is
reasonably well established however the area of signal processing still offers substantial gains.

A reasonable solution to the sensing problem would be a system that can estimate the remnant layer thickness over the range close to zero to 50 cm or further. Furthermore, it is advantageous if the sensor processes the data in near real-time, automatically generates the layer thickness estimate without the need for interpretation and is capable of sensing multiple layers. The acquisition rate of GPR data is very fast, therefore the near real-time requirement is predominantly governed by signal processing. Solutions for the automatic processing of GPR data for layer thickness estimation have been proposed in the literature however system induced clutter impose a minimum detectable thickness limit. Finally, GPR is suitable for imaging multi-layered environments. Therefore it is concluded that GPR is a prime candidate for the task of remnant layer thickness estimation for horizon control sensing. However the minimum detectable thickness limitation imposed by system induced clutter must be addressed.

Therefore the specific research objectives of this dissertation are as follows.

1. To develop a signal processing method to estimate the layer thickness of a thin top layer with GPR in the presence of system clutter.

2. To develop signal processing methods to enhance the thin layer thickness estimator to improve robustness.

3. To investigate the effects of environmental parameters such as surface dielectric constant, conductivity, and GPR system noise on the developed methods.

4. To test the developed methods on ground structures similar to that expected in an underground coal mine.

This chapter has introduced the need for coal-rock interface sensing for underground coal mining operations and the need for thin layer thickness estimation and interface detection. The next chapter presents all matters relating to the collection of both synthetic and real GPR data used in this dissertation.
Chapter 3

Experimental Data Collection

3.1 Introduction

The typical underground coal mine is a complex and hazardous environment. Some of the difficulties in acquiring experimental data from an underground coal mine are: using electrical equipment underground; obtaining intrinsically safe (IS) approval for equipment; unknown conditions and drilling to obtain ground truth information. A twofold approach has been adopted to alleviate these practical difficulties.

Synthetic data used for algorithm testing was generated using the finite-difference time-domain (FDTD) method. Data for validation purposes was acquired using a low power impulse GPR system and a testbed with coal, shale and clay layers. The testbed was constructed to represent a test coal seam. The FDTD parameters were configured to closely match the real data acquired from the GPR system and testbed.

The chapter begins with the design of the experimental testbed, ground truth measurements of layer thicknesses and the measured electromagnetic properties of the target media. The GPR system used to acquire real data is then described. Finally, the method implemented to generate synthetic data based on parameters of the testbed and GPR system is discussed.
3.2 Experimental Testbed

3.2.1 Design

A testbed was designed and constructed to validate the proposed processing methods with real GPR data acquired from media similar to that found in an underground coal mine. The testbed has dimensions of $2.4 \times 2.25 \times 0.8$ m deep and the frame is timber. The testbed depth was chosen such that the floor is out of range of the GPR system used for the experiments.

The horizontal layout of the testbed was divided into 15 regions where each region has either two or three layers. The bottom layer in all regions consists of wet clay with volumetric water content of 15%. The top layer for all regions except 1 and 5 is unconsolidated coal representing the product which would be extracted during the mining process. Regions 1–4 and 11–15 have a thin layer of shale between the bottom clay layer and top coal layer. The top layer for regions 1 and 5 is shale and clay respectively, representing the case when the coal layer has been mined right down to the surrounding strata. As it is impractical to obtain high quality coal and low quality coal, the unconsolidated coal and shale were considered reasonable substitutes and they were readily available from a local quarry. All layers were separated by thin plastic sheets to prevent evaporation and keep the volumetric water content constant over time. Figure 3.1 shows a photo of the testbed. Figures 3.2 and 3.3 show the plan view and side profile of the testbed respectively.
3.2. Experimental Testbed

Figure 3.1: Photo of the GPR testbed designed and constructed for ground truth evaluation.

Figure 3.2: Plan view of the GPR testbed. The testbed has 14 stepped regions and a linear ramp. The layers consist of coal, shale and clay of varying thickness. The bottom layer in all regions is clay. There is no shale in regions 5–10 as these are for investigating coal-clay interfaces.
Figure 3.3: Side profile of the layers for all regions of the GPR testbed. The values shown correspond to the average thicknesses for a given region.
3.2.2 Ground Truth for Validation

The ground truth layer thicknesses were obtained by measuring the interface profiles for each layer during construction. Measurements were taken at spatial intervals of 20 cm intervals in both $x$ and $y$ directions and at region boundary corners using a tape measure with 1 mm resolution, as shown in Figure 3.4. The circles denote the measurement points 20 cm apart whereas the diamonds represent the region boundary corners. Depth measurements of all the testbed layers and regions were taken as these points during the construction phase of the testbed. The coal layer thickness range is 0–37 cm while the shale layer thickness varies between 2–7 cm in the shallow sections and up to 19 cm in the deepest section (region 14).

A strip of the testbed was also measured at 50 mm spacing along a single traverse line. The side profile of this area is shown in Figure 3.5. Figure 3.5(a) shows the surface topography of the strip and Figure 3.5(b) shows the layer depths adjusted to a flat surface. The measured thickness of the coal and shale layers are shown in Appendix A.

![Figure 3.4: Plan view of the GPR testbed detailing the ground truth measurement points. Region 1 is shown in the top left along with the adjacent regions.](image-url)
Figure 3.5: Plots showing side profile of the testbed area measured at 50 mm spacing with (a) true surface contour and (b) profile referenced to a flat surface.
3.2.3 Measurement of Electromagnetic Properties

The electromagnetic properties of media can be described with the real ($\varepsilon'_r$) and imaginary relative permittivity ($\varepsilon''_r$). Other parameters derived from the complex permittivity are the loss tangent ($\tan \delta$) and conductivity ($\sigma$) \[92\]. Media is dispersive when these parameters are frequency dependant and results in frequency dependant wave-propagation velocity and attenuation.

The electromagnetic properties of the coal, shale and clay were measured using an HP8752A network analyzer with an open-ended coaxial surface probe. The frequency range measured was 100 MHz–3 GHz. The average values over the full measured range are shown in Figure 3.6. The minimum, maximum and average (in brackets) values over the antenna bandwidth for the real and imaginary relative permittivity, loss tangent and conductivity (in mS/m) are shown in Table 3.1.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coal</th>
<th>Shale</th>
<th>Clay</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon'_r$</td>
<td>3.8–5.3 (4.4)</td>
<td>3.1–12.7 (9.1)</td>
<td>22.7–27.2 (24.9)</td>
</tr>
<tr>
<td>$\varepsilon''_r$</td>
<td>0.1–0.5 (0.26)</td>
<td>0.2–4.5 (2.1)</td>
<td>5.2–9.9 (6.9)</td>
</tr>
<tr>
<td>$\tan \delta$</td>
<td>0.02–0.1 (0.05)</td>
<td>0.1–0.4 (0.2)</td>
<td>0.2–0.4 (0.3)</td>
</tr>
<tr>
<td>$\sigma$ mS/m</td>
<td>14.3–25.6 (20.5)</td>
<td>133.1–176.4 (154.6)</td>
<td>438.6–602.1 (517.6)</td>
</tr>
</tbody>
</table>

Table 3.1: Range of measured electrical parameters of the testbed layers. The average values over the GPR antenna bandwidth are shown in brackets.

3.3 GPR Hardware

3.3.1 Description

The GPR system used for the experimental work is a low power impulse GPR custom built by CSIRO for short range coal mining applications \[70\]. It has a pulse duration of 1–2 ns, dynamic range $> 50$ dB, and has been adapted for use in an underground coal mine. The GPR has a free-space operating range of 60 cm.

The bi-static antenna module consists of two shielded bow-tie antennas with resistive loading on a dielectric half space. The transmit and receive antennas have a separation distance of 5 cm and centre frequency of 1.4 GHz in air with
Chapter 3. Experimental Data Collection

(b) Imaginary relative permittivity ($\varepsilon''_r$)

(c) Loss tangent ($\tan \delta$)

(d) Electrical conductivity ($\sigma$)

Figure 3.6: Measured electrical parameters of coal, shale and clay from testbed from 100 MHz–3 GHz.

bandwidth of 1.1 GHz. The flare angle of the bow-tie antennas is 90° and the maximum dimension is 7 cm resulting in a near-field far-field threshold of 6 cm [92].

The GPR electronics utilises sequential sampling [30] to down-sample the high frequency signal for sampling by a low speed analog to digital converter. In this method, one sample is acquired from a transmitted pulse at a time delay of $n\Delta t$ after the start of the trace, where $n$ is the sample number and $\Delta t$ is the time delay. This process is repeated until an entire waveform has been acquired, at which point a new trace is commenced. For this system, the equivalent sampling rate is 111.3 GHz which results in a $\Delta t$ of approximately 8.98 ps. The downsampled data is acquired via a laptop computer with a 12-bit PCMCIA data acquisition card configured for 12-bit uniform bipolar sampling over the range
3.3. **GPR Hardware**

±2.5 V. Each waveform consists of 512 samples sampled at the GPR system pulse repetition frequency (PRF) of 28.8 kHz. Figure 3.7 shows the GPR system used for the experiments. Figure 3.8 shows a typical raw waveform acquired from this GPR system. A noise analysis of the GPR system is presented in Appendix B.

![Figure 3.7](image1)

(a)

![Figure 3.7](image2)

(b)

Figure 3.7: Photos of (a) the GPR electronics and (b) the GPR system with laptop computer used for data acquisition.
Figure 3.8: Typical raw A-scan from GPR system
3.4 Synthetic Data Generation

3.4.1 The Finite-Difference Time-Domain Method

The finite-difference time-domain (FDTD) method is a powerful computational electromagnetics modelling tool based on the discretization of Maxwell’s time domain curl equations. First proposed by Yee [126] in 1966, it is very popular due to its ability to model complicated and inhomogeneous structures.

The FDTD method was implemented in two-dimensions (2D) to generate synthetic data for testing. The second order absorbing boundary condition developed by Mur [69] was implemented to reduce artificial reflections from the computational space boundaries. The FDTD update equations and absorbing boundary conditions implemented for the synthetic data generation are shown in Appendix C.

The transmitting antenna was modelled as a small dipole polarized in the $z$ direction by a single cell. Similarly, the receiver was modelled as a small dipole that samples the $z$ component of the electric field. The time increment was chosen as 8.98 ps, which is the period of the equivalent sampling rate of the test GPR system presented in Section 3.3. The number of time steps was 500 and the cell spacing in both the $x$ and $y$ directions ($\Delta x = \Delta y$) was 0.5 cm per cell. The number of cells in both the $x$ and $y$ directions was 250 representing a synthetic area of 1.25 m by 1.25 m.

The electromagnetic parameters of inhomogeneous materials are often frequency dependant. This property is the cause of dispersion as the energy propagates through the layers. Dispersion was modelled in the FDTD grid using the Debye relaxation model and the recursive convolution method developed by Luebbers et al. [58] but adapted into 2D form as proposed by Bourgeois [11]. The Debye model for frequency dependant relative permittivity is:

$$\varepsilon_r(\omega) = \varepsilon'_r(\omega) - j\varepsilon''_r(\omega) = \varepsilon_{\infty} + \frac{\varepsilon_{rs} - \varepsilon_{\infty}}{1 + j\omega\tau}$$

(3.1)

where $\varepsilon_{rs}$ is the static frequency relative permittivity, $\varepsilon_{\infty}$ is the infinite frequency relative permittivity, $\tau$ is the relaxation time and $\omega$ is radial frequency in radians.
The real relative permittivity ($\varepsilon'_r$) and conductivity ($\sigma$) electromagnetic parameters described in Section 3.2.3 are related to the Debye model parameters as follows:

$$\varepsilon'_r(\omega) = \varepsilon_{r\infty} + \frac{\varepsilon_{rs} - \varepsilon_{r\infty}}{1 + (\omega \tau)^2}$$  \hspace{1cm} (3.2)

$$\sigma(\omega) = \sigma_o + \omega \varepsilon''_r(\omega) \varepsilon_o = \sigma_o + \frac{\varepsilon_{rs} - \varepsilon_{r\infty}}{1 + (\omega \tau)^2} \varepsilon_o \omega^2 \tau$$  \hspace{1cm} (3.3)

where $\sigma_o$ is the low frequency conductivity.

The values for the Debye model parameters chosen for the FDTD grid are shown in Table 3.2 and were obtained by fitting the measured testbed material properties shown in Figures 3.6(a) and 3.6(d) to equations (3.2) and (3.3).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Coal</th>
<th>Shale</th>
<th>Clay</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{rs}$</td>
<td>4.7</td>
<td>10.7</td>
<td>26</td>
</tr>
<tr>
<td>$\varepsilon_{r\infty}$</td>
<td>4.21</td>
<td>8.6</td>
<td>22</td>
</tr>
<tr>
<td>$\tau$ (ps)</td>
<td>140</td>
<td>180</td>
<td>70</td>
</tr>
<tr>
<td>$\sigma_o$ (mS/m)</td>
<td>2</td>
<td>85</td>
<td>370</td>
</tr>
</tbody>
</table>

Table 3.2: Debye relaxation parameters chosen for FDTD simulations.

### 3.4.2 Modified FDTD: Real Source Signal

The major problem with detecting near-surface targets using ground-coupled antennas arises when target reflections are masked by the antenna ringing and crosstalk. To completely model the ringing, a 3D FDTD computational space must be implemented that models the major components of the antenna structure. As a 3D FDTD implementation is significantly more computationally expensive compared with 2D, an alternative was investigated such that the antenna ringing can be represented in the synthetic data.

From investigation of the FDTD output, the sampled electric field is approximately proportional to the temporal derivative of the source function. A polynomial curve was fitted to an ensemble averaged free-space scan acquired from the GPR system. This polynomial was integrated and incorporated as the source function. The mathematical details of this process are described below.
The real GPR data from an averaged free-space scan was represented by a polynomial [51]

\[ r(t) = \sum_{k=0}^{M} a_k t^k \quad t = 0, 1, \ldots, N - 1 \]  

(3.4)

where \( M \leq N - 1 \) is the degree of the polynomial, \( N \) is the number of samples in the radar data, \( a \) are the polynomial coefficients and \( t = 0, 1, \ldots, N - 1 \) is the time of each sample. The polynomial degree implemented was 30 which provided good smoothing without losing features. ¹ The expanded form of the polynomial in (3.4) is:

\[
\begin{align*}
    r(t_0) &= a_0 + a_1 t_0^1 + a_2 t_0^2 + \cdots + a_M t_0^M \\
    r(t_1) &= a_0 + a_1 t_1^1 + a_2 t_1^2 + \cdots + a_M t_1^M \\
    \vdots \\
    r(t_{N-1}) &= a_0 + a_1 t_{N-1}^1 + a_2 t_{N-1}^2 + \cdots + a_M t_{N-1}^M
\end{align*}
\]  

(3.5)

where \( t_0 = 0, t_1 = 1, \ldots, t_{N-1} = N - 1 \). Equation (3.5) can be represented in matrix form as:

\[
\begin{bmatrix}
    r(t_0) \\
    r(t_1) \\
    \vdots \\
    r(t_{N-1})
\end{bmatrix} =
\begin{bmatrix}
    1 & t_0 & t_0^2 & \cdots & t_0^M \\
    1 & t_1 & t_1^2 & \cdots & t_1^M \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    1 & t_{N-1} & t_{N-1}^2 & \cdots & t_{N-1}^M
\end{bmatrix}
\begin{bmatrix}
    a_0 \\
    a_1 \\
    \vdots \\
    a_M
\end{bmatrix}
\]  

(3.6)

This can be re-expressed using matrix notation as

\[
r = Xa
\]  

(3.7)

where \( X \) is a \([N \times (M + 1)]\) matrix where each column is the \( t^k \) time values, \( r \) is a \([N \times 1]\) column vector of the observed GPR data and \( a \) is the \([(M + 1) \times 1]\) column vector of the polynomial coefficients to be estimated. Multiplying through by the transpose of matrix \( X \), i.e., \( X^T \), results in the familiar normal equations [48], which can be solved using

\[
X^T r = X^T X a
\]  

(3.8)

¹Orthogonalised polynomial basis sets or splines may be used under conditions when inversion problems arise.
\[ a = (X^T X)^{-1} X^T r \] 

(3.9)

Numerically stable and computationally efficient algorithms exist to solve for the polynomial coefficients \( a \) in a least squares sense.

The polynomial was integrated and evaluated to become the source signal. The free-space GPR scan and FDTD scan using the integrated free space scan as the source are shown in Figure 3.9 for both ground-coupled and air-coupled antenna configurations. The figure shows that the synthetic FDTD data using the integrated form of the real free space scan closely resembles the measured GPR data. This is, therefore, an acceptable method of generating synthetic GPR data that models system clutter such as the crosstalk and antenna ringing.

3.5 Summary

This chapter has described the source of GPR data that was used for testing and validating the processing techniques proposed in this dissertation. Firstly, the design and construction of a testbed to that enabled the acquisition of real GPR data under known conditions was presented. The measurements of the electromagnetic parameters of coal, shale and clay samples used in the testbed were provided. The GPR system used to acquire real GPR data from the testbed was described. Finally, the popular method of generating synthetic GPR data, the FDTD method, was presented along with the specific FDTD parameters used to generate synthetic GPR data that resembles the real data acquired from the GPR system. Next, a method to model clutter components such as the antenna ringing and crosstalk using the FDTD method without the need for a complete 3D model was proposed.
3.5. Summary

(a) Ground-coupled antennas

(b) Air-coupled antennas

Figure 3.9: Real and synthetic GPR traces for ground and air-coupled antennas. The ground-coupled antennas are in contact with a coal surface.
Chapter 4

Early-Time Feature Vector for Near-Surface Processing

4.1 Introduction

This chapter presents a novel feature vector that characterises the early-time signal of 1D GPR data. The early-time signal of the GPR data is the beginning segment where the crosstalk and antenna ringing dominates the received GPR trace. The feature vector is based on an approach derived from the bispectrum that was developed primarily for pattern recognition applications.

As described in Section 2.3.2, the important practical problems associated with the use of classical analysis methods to estimate thin layer thickness strongly motivate the investigation into alternate processing techniques. To this end, the use of pattern recognition techniques for this problem is proposed.

To begin the chapter, a brief introduction to the field of pattern recognition and a review of where pattern recognition has been applied to GPR is provided. A review of the signal processing domain adopted for the feature vector, higher order spectra and the bispectrum, is then presented. The proposed feature vector is then presented along with an analysis which includes feature values estimated from synthetic GPR data of various known sub-surface configurations applicable to the thin layer thickness estimation task. Finally, the statistical properties of the proposed feature vector and noise reduction capabilities with averaging are
4.2 Pattern Recognition

Pattern recognition is the act of taking raw data and making an action based on the category of the pattern [35]. A typical pattern recognition system can be separated into five main components – sensing, segmentation, feature extraction, classification and post-processing [35]. A flow chart of these components is shown in Figure 4.1.

![Diagram of a typical pattern recognition system](image)

Figure 4.1: A typical pattern recognition system consists of five main components. The bold components, segmentation and feature extraction, are addressed in this chapter.

The input is *sensed* by a transducer of some kind and converted into a signal. The sensor is the GPR system in the current investigation. The received signal is *segmented* to isolate sensed objects from background and other objects that are not of interest. A *feature extractor* measures certain properties of the segmented
signal which is then allocated to a category by a classifier. The data is post-processed to account for other considerations and then the resulting action of the pattern recognition system is then determined.

The sensing stage for the pattern recognition system proposed in this dissertation was described in Chapter 3. The components covered in this chapter are segmentation and feature extraction. The classification and post-processing stages are described in Chapters 5, 6 and 7 respectively.

4.2.1 Pattern Recognition for GPR

There are many reported cases where pattern recognition has been applied to GPR processing tasks. In general, these works have been focused on the detection and/or classification of two separate target classes – (a) military based such as landmines and unexploded ordnance (UXO), and (b) utility based such as buried pipes, cables and voids.


Utility based investigations include the following. Kleinmann et al. [52] used the Hough transform to detect and identify point reflectors and texture analysis to determine soil structure. Bishop et al. [9] used pre-processing algorithms and
a neural network to detect sub-surface voids. Bisessur and Naguib [10] used the overlap weighted discrete Fourier transform (DFT) and a neural network classifier to detect underground pipes. Al-Nuaimy [2] presented novel processing techniques for the automatic detection and interpretation of GPR data from targets such as landmines, UXO, cables, pipes and extended plumes of hydrocarbon-contaminated soil. Al-Nuaimy et al. [3] used features derived from the power spectrum for underground pipe detection and developed a novel adaptive 3D Hough transform algorithm for the automatic mapping of linear targets such as buried pipes and cables [4]. Shihab et al. [98] presented a comparison of three techniques for automatic GPR data segmentation to discriminate between subsurface targets. Shihab and Al-Nuaimy [95] proposed a method for automatic fusion of multiple B-scans to determine the geometry of buried pipes [95] and presented a model for the hyperbolic signature of buried cylindrical targets [96]. They also presented a novel technique for automatic segmentation and characterisation of B-scan images [97]. This latest approach utilised a neural network classifier to identify hyperbolic signatures of buried utility pipes.

Other investigations include the following. De Pasquale and Pinelli [33] investigated the use of pattern recognition to recognise different soil types. Koppenjan et al. [54] proposed the use of singular-value decomposition for object recognition in a forensic application.

Unlike all of the previous works in this review, pattern recognition as in the framework described in Figure 4.1 has been seldom been applied to the detection and thickness estimation of planar targets. It is, therefore, valuable to explore pattern recognition for the thin layer thickness estimation task.

It has been shown by Al-Nuaimy et al. [4] that the power spectrum is a useful feature vector for a pattern recognition based underground pipe detection system. It has also been reported that the Fourier phase contains information that represents the shape of a signal [78]. It is well known however that the power spectrum disregards Fourier phase information. Therefore a signal processing domain that retains both magnitude and phase information is a prime candidate for a feature vector. One domain that retains phase information is higher order spectra (HOS). A review of HOS is presented in the next section.
4.3 Higher Order Spectra

Higher order spectra (HOS), also called polyspectra, is an area that has received significant attention by the signal processing community in the past 25 years [110]. HOS are defined in terms of higher order statistics [75] (moments and cumulants) and can be applied to both stochastic and deterministic signals. As such, there are two categories of polyspectra – moment spectra and cumulant spectra.

In general, there are four motivations for using HOS in signal processing. The first motivation is to suppress Gaussian noise [75]. This arises from the observation that for Gaussian processes, cumulant spectra of orders greater than two are zero. Another observation is that for non-Gaussian processes with symmetric probability density functions, both cumulant and moment spectra with odd orders are also zero.

The second motivation for using HOS is to reconstruct the phase and magnitude response of signals or systems [75]. It is well known that second order processes such as the power spectrum and autocorrelation function do not retain phase information, and hence cannot be used to reconstruct the phase response of a signal unless it is minimum phase. However HOS with order greater than two preserve the true phase character of a signal and hence can be used to reconstruct both magnitude and phase information for non-minimum phase processes. The third motivation is to detect and characterise non-linearities in time series data [75]. The fourth motivation is that second order statistics such as the power spectrum do not, in general, provide a complete characterisation of non-Gaussian data.

The power spectrum is a special case of HOS and has an order of two. Apart from the power spectrum, the most widely used orders of HOS are the third and fourth orders, the bispectrum and trispectrum.

Polyspectra has been used for a wide variety of applications such as system identification [60, 109, 119], pattern recognition [18, 21, 115], geophysics [41], oceanography [36], radar [46, 130] and sonar processing [43, 112]. There are several bibliographies of higher order statistics and spectra publications that have been published in the literature [73, 110, 114]. There have also been some key
tutorial papers published by Nikias and Raghuveer [76] and Mendel [61].

### 4.3.1 Moment and Cumulant Statistics

For a set of \( n \) real random variables, \( \{x_1, x_2, \ldots, x_n\} \), the joint moment function of order \( r = k_1 + k_2 + \cdots + k_n \) is given by [76]

\[
m_{k_1 k_2 \cdots k_n} \triangleq E \{x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}\}
\]

\[
= (-j)^r \frac{\partial^r \Phi (\omega_1, \omega_2, \ldots, \omega_n)}{\partial \omega_1^{k_1} \partial \omega_2^{k_2} \cdots \partial \omega_n^{k_n}} \bigg|_{\omega_1=\omega_2=\cdots=\omega_n=0} \tag{4.1}
\]

where

\[
\Phi (\omega_1, \omega_2, \ldots, \omega_n) \triangleq E \{\exp (j (\omega_1 x_1 + \omega_2 x_2 + \cdots + \omega_n x_n))\} \quad \tag{4.2}
\]

is the joint characteristic function and \( E \{\cdot\} \) denotes the expectation operation.

For the same set of \( n \) real random variables, the joint cumulant function is [76]

\[
c_{k_1 k_2 \cdots k_n} \triangleq \text{Cum} [X(k)X(k + \tau_1) \ldots X(k + \tau_{n-1})] \tag{4.3}
\]

\[
= (-j)^r \frac{\partial^r \ln \Phi (\omega_1, \omega_2, \ldots, \omega_n)}{\partial \omega_1^{k_1} \partial \omega_2^{k_2} \cdots \partial \omega_n^{k_n}} \bigg|_{\omega_1=\omega_2=\cdots=\omega_n=0} \tag{4.4}
\]

For real discrete-time stationary random process \( X(k) \) with \( k = 0, \pm 1, \pm 2, \ldots \) and its moments up to \( n \) exist, the \( n \)th-order moment function can be defined by [74]

\[
m_n^x (\tau_1, \tau_2, \ldots, \tau_{n-1}) \triangleq E \{X(\tau_1) X(\tau_2 + \tau_1) \ldots X(\tau_{n-1} + \tau_{n-2})\} \tag{4.5}
\]

where \( \tau_i = 0, \pm 1, \pm 2, \ldots \) for all values of \( i \). This moment function only depends on the time differences \( \tau_1, \tau_2, \ldots, \tau_{n-1} \).

Similarly, the \( n \)th-order cumulant sequence of a real stationary random process can be defined by [75]

\[
c_n^x (\tau_1, \tau_2, \ldots, \tau_{n-1}) \triangleq \text{Cum} [X(\tau_1) X(\tau_2 + \tau_1) \ldots X(\tau_{n-1} + \tau_{n-2})] \tag{4.6}
\]

### 4.3.2 Moment and Cumulant Spectra

Polyspectra are defined as the Fourier transform of moment and cumulant statistics. Therefore, the moment spectra of a real stationary random process with \( n \)th-order moment function \( m_n^x (\tau_1, \tau_2, \ldots, \tau_{n-1}) \) can be defined as
4.3. Higher Order Spectra

\[ M^x_n (\omega_1, \omega_2, \ldots, \omega_{n-1}) \triangleq \sum_{\tau_1 = -\infty}^{+\infty} \cdots \sum_{\tau_{n-1} = -\infty}^{+\infty} m^x_n (\tau_1, \tau_2, \ldots, \tau_{n-1}) \]
\[ \cdot \exp \{ -j (\omega_1 \tau_1 + \omega_2 \tau_2 + \cdots + \omega_{n-1} \tau_{n-1}) \} \] (4.8)

where \( |\omega_i| \leq \pi \) for \( i = 1, 2, \ldots, n-1 \) and \( |\omega_1 + \omega_2 + \cdots + \omega_{n-1}| \leq \pi \).

For a real strictly stationary random process \( X(k) \) with \( n \)-th order cumulant sequence \( c_n (\tau_1, \tau_2, \ldots, \tau_{n-1}) \), the \( n \)-th order cumulant spectrum is defined as

\[ C^c_n (\omega_1, \omega_2, \ldots, \omega_{n-1}) \triangleq \sum_{\tau_1 = -\infty}^{+\infty} \cdots \sum_{\tau_{n-1} = -\infty}^{+\infty} c_n (\tau_1, \tau_2, \ldots, \tau_{n-1}) \]
\[ \cdot \exp \{ -j (\omega_1 \tau_1 + \omega_2 \tau_2 + \cdots + \omega_{n-1} \tau_{n-1}) \} \] (4.9)

where \( |\omega_i| \leq \pi \) for \( i = 1, 2, \ldots, n-1 \) and \( |\omega_1 + \omega_2 + \cdots + \omega_{n-1}| \leq \pi \).

The conditions for existence of the moment and cumulant spectra are that the moment and cumulant statistic sequences must be absolutely summable [75].

### 4.3.3 The Bispectrum

The bispectrum is the third order polyspectrum and can be defined for both stochastic and deterministic signals using moment and cumulant spectra. It is common practice, however, to model stochastic signals using cumulant spectra and deterministic signals with moment spectra. The key motivating factor for using cumulant spectra for stochastic signal processing is based on the fact that for Gaussian processes, cumulants with order greater than two are identically zero [75]. Moment spectra are used for analysing deterministic signals because there is no clear advantage by using cumulant spectra for this case [75]. As one trace of the GPR signal can be modelled as a deterministic but unknown transient signal plus noise, the bispectrum defined using moment spectra for deterministic energy signals will be developed here.

The \( n \)-th order moment function of energy signal \( x(k) \) with \( k = 0, \pm 1, \pm 2, \ldots \) can be defined by [75]

\[ m^x_n (\tau_1, \ldots, \tau_{n-1}) \triangleq \sum_{k = -\infty}^{+\infty} x(k)x(k + \tau_1) \cdots x(k + \tau_{n-1}) \] (4.10)
where \( \tau_i = 0, \pm 1, \pm 2, \ldots \) for all values of \( i \).

A special case of the previous equation is the third order moment function, also called the \textit{triple correlation} and is

\[
m_{3}^{x}(\tau_1, \tau_2) = \sum_{k=-\infty}^{+\infty} x(k)x(k+\tau_1)x(k+\tau_2) \tag{4.11}
\]

The moment spectra of an energy signal, in the form of the Fourier transform of the moment function, is shown as equation (4.8). The moment spectra can also be expressed in terms of the Fourier transform of the energy signal as

\[
M_{n}^{x}(\omega_1, \omega_2, \ldots, \omega_{n-1}) = X(\omega_1)X(\omega_2)\cdots X(\omega_{n-1})X^{*}(\omega_1 + \omega_2 + \cdots + \omega_{n-1}) \tag{4.12}
\]

As the bispectrum is a special case of polyspectra with order \( n = 3 \), the bispectrum can be defined as the third-order moment spectrum

\[
B(\omega_1, \omega_2) = M_{3}^{x}(\omega_1, \omega_2) = X(\omega_1)X(\omega_2)X^{*}(\omega_1 + \omega_2) \tag{4.13}
\]

where \( |\omega_1| \leq \pi, |\omega_2| \leq \pi \) and \( |\omega_1 + \omega_2| \leq \pi \).

Both moment and cumulant statistics have important symmetry properties. The triple correlation domain in (4.11) can be segmented into six sectors as shown in Figure 4.2(a). The first sector is an infinite wedge bounded by the lines \( \tau_1 = 0, \tau_1 = \tau_2 \) and \( \tau_1, \tau_2 \geq 0 \). The entire third-order moment function (all six segments) can be computed from knowledge of just one of the segments.

These symmetry properties are also extended to the bispectrum. The bispectrum can be segmented into twelve sectors as shown in Figure 4.2(b). The boundary surrounding the twelve sectors in Figure 4.2(b) results from the frequency limits imposed on \( \omega_1 \) and \( \omega_2 \) in (4.13). Due to the symmetry of the bispectrum, the triangle labeled 1 in Figure 4.2(b) is commonly referred to as the non-redundant sector. As the bispectrum is repeated in the other sectors, the entire bispectrum can be estimated from knowledge of just one sector.
4.3. Higher Order Spectra

\[ \tau_2 \omega_2 = \tau_1 \]

\[ \tau_1 = \tau_2, \omega_1 = \omega_2 \]

(a) Third order moment

(b) Bispectrum

Figure 4.2: Symmetry properties of (a) the third order moments and (b) the bispectrum.

4.3.4 Bispectrum Estimation

There have been a variety of approaches proposed in the literature for estimating the bispectrum of a signal. The two main approaches are the conventional and parametric types [84]. Parametric type approaches are based on auto-regressive (AR), moving average (MA), auto-regressive moving average (ARMA) and Volterra models [74]. The conventional approaches are based on techniques developed for power spectrum estimation and include three classes – indirect, direct and complex demodulates. There are advantages and disadvantages for choosing a certain class such as computational complexity, statistical properties and resolution.

The bispectrum is a multi-dimensional function of two frequencies. Hence its estimation is a computationally complex process. One method proposed to reduce the computational complexity of bispectrum estimation was proposed by Bessios and Nikias [7]. This method involves computing the bispectrum along radial slices in the first sector in Figure 4.2(b). This approach is a more efficient method of computing bispectrum estimates than conventional methods. Another approach was proposed by Chandran and Elgar [21] where the bispectrum is estimated radially along lines in the first sector, interpolated bilinearly in the bi-frequency plane and then integrated. This approach was developed as a feature
vector to satisfy certain invariance properties for pattern recognition applications. This method is central to the proposed feature vector and is discussed in the next section.

### 4.3.5 The Integrated Bispectrum

The integrated bispectrum developed by Chandran and Elgar [19, 20, 21] has been used in a wide range of pattern recognition applications such as image processing [25, 57, 113], sea-mine detection [22, 23], speaker recognition [24] and virus identification [77]. The approach has also been used to detect and classify buried landmines with two-dimensional GPR data [6].

The integrated bispectrum is computed by integrating radially along slices in the bi-frequency plane. This approach was motivated by the need to obtain certain invariant properties that are attractive to pattern recognition tasks. The integrated bispectrum presented here is not the same as the integrated bispectrum method proposed in [119, 120, 124] for system identification and [116, 117, 118] for the detection of non-Gaussian signals. These works defined the integrated bispectrum as the cross-spectrum between the signal and its square. That method and the one used in this dissertation are not the same.

It could be considered that the work by Ye and Tugnait [125] to estimate time delay using the integrated polyspectrum may be useful for the current investigation. In this work, they estimated the time delay between two or more receiver sensors. The GPR system used in this application has only one receiver sensor. Modifying the GPR hardware to have multiple receivers poses additional unnecessary challenges and is impractical for the application described in the dissertation due to the harsh operating environment of an underground coal mine.

The references of the integrated bispectrum in the remainder of this dissertation specifically relate to the method proposed by Chandran and Elgar [21].

The bispectrum, $B(f_1, f_2)$, of a deterministic discrete-time transient signal, $x(n)$, can be defined as

$$B(f_1, f_2) = X(f_1)X(f_2)X^*(f_1 + f_2) \tag{4.14}$$
where $X(f)$ is the discrete-time Fourier transform of $x(n)$, $f$ is frequency normalised by one half the sampling rate and $\ast$ is the complex conjugate operator. As per the symmetry properties described in Section 4.3.3, the bispectrum is defined in the triangular region, $0 \leq f_2 \leq f_1 \leq f_1 + f_2 \leq 1$, provided there is no bispectral aliasing [21]. Bispectral aliasing is avoided by sampling the continuous time signal by at least three times the highest frequency component [108].

To obtain a feature that is more immune to noise and robust to small frequency changes, the bispectrum is integrated [21]

$$I(a) = \int_{f_1=0^+}^{1/(1+a)} B(f_1, af_1) df_1$$

where $I(a)$ is the integrated bispectrum, $a$ is the slope of the radial line of integration and satisfies $0 < a \leq 1$, and $j = \sqrt{-1}$. To obtain multiple features for a higher dimensional feature vector, the values of $a$ are chosen to be evenly spread between 0 and 1. If only one feature value is required, $a$ is usually chosen as 1. In practice, the value of $a$ which provides the most discrimination between the classes is the best choice for a single feature. If multiple features are used, the problem of choosing the optimal combination requires a measure of effective discrimination or error statistics. Figure 4.3 shows the radial lines of integration for $a$ parameters of 0.25, 0.5, 0.75 and 1.

![Radial lines of integration](image)

Figure 4.3: The bispectrum is integrated along lines with slope $a$ chosen to be spread evenly between 0 and 1. When four parameters are computed, the values for $a$ are 0.25, 0.5, 0.75 and 1. This triangular region is sector 1 in Figure 4.2(b).

The phase of the integrated bispectrum satisfies several invariant properties. These are translation, DC level, amplification, and scale invariance [21]. Even
though these bispectral-based features cannot be associated with any physical electromagnetic scattering phenomena, the features represent signal shape information in a useful way offering discrimination between certain classes. The phase of the integrated bispectrum is is computed using the following.

\[ P(a) = \angle I(a) = \arctan \left( \frac{I_i(a)}{I_r(a)} \right) \]  

where \( I_r(a) \) and \( I_i(a) \) are the real and imaginary components of the integrated bispectrum.

### 4.4 Proposed Early-Time Feature Vector

The previous section presented a review of the integrated bispectrum. This section introduces the novel feature vector to characterise the early-time component of the GPR data. The proposed feature vector comprises of two important processing phases – segmentation and feature extraction. These are discussed in the following sections.

#### 4.4.1 Segmentation

The aim of segmentation is to isolate signal regions that either degrade performance or are not of interest to the current task. The region of interest for thin layer processing is both system and application dependant. For the GPR system and antenna module used in this study, the dominant part of the antenna ringdown and crosstalk is limited to approximately the first 200 samples.

This region where the signal is dominated by the clutter components of crosstalk and antenna ringing is segmented using a window function. Two window types were considered for this task – the rectangular and Hamming windows. The Hamming window was investigated as it provides a good tradeoff between mainlobe resolution and sidelobe attenuation [42]. The rectangular window was also considered as the shape of the segmented signal remains unchanged. Figure 4.4 shows an example of each of the windows.

The starting point of the window is chosen at the time-zero point. Time-zero is the point in the GPR signal to which propagation delays are referenced for layer
4.4. Proposed Early-Time Feature Vector

Figure 4.4: Window functions used in the feature extraction phase.

Figure 4.4: Window functions used in the feature extraction phase.

Varying the window length provides an element of control of target depth threshold. A shorter window length selects the region that is modified by closer targets and hence has the effect of better resolution for closer targets. A longer window however contains more signal and hence the feature values can be dependent upon the presence of targets not as close to the surface. Ultimately the best choice of window length is what gives the best cluster separation between classes for the desired target depth threshold.

Figure 4.5 shows a synthetic A-scan overlayed with the same signal windowed with varying window lengths from 100 to 200 samples. This figure indicates how much of the early-time signal is included in the feature vector for a given window length.

4.4.2 Feature Extraction

The integrated bispectral phase parameter presented in Section 4.3.5 is a useful feature for discriminating between signals based on their shape. Additional information however is available in the magnitude of the integrated bispectrum.
Chapter 4. Early-Time Feature Vector for Near-Surface Processing

This provides additional discrimination between different classes of the early-time GPR data. The main motivations for using the integrated bispectrum are (a) both magnitude and phase information is retained compared with the power spectrum that just retains the magnitude, (b) robustness to small frequency changes resulting from the integration (scale invariance), (c) computationally efficient approach compared with conventional bispectrum estimation methods and (d) noise reduction due to integrating. The early-time feature vector consists of the magnitude and phase of the integrated bispectrum and can be expressed as

$$\bar{x} = \begin{bmatrix} |I(a)| & P(a) \end{bmatrix}$$ (4.17)

As reported by Swami [108], there are certain factors that must be considered when estimating the bispectrum. Consider a sampled signal that is bandlimited over the range \([B_1, B_2]\) and has a sampling rate that satisfies the Nyquist criterion (ie, \(f_s \geq 2B_2\)). The first consideration relates to the signal bandwidth. If the bandwidth is less than one octave (\(2B_1 \geq B_2\)), the bispectrum over the triangular region in Figure 4.3 (sector 1 in Figure 4.2(b)) is identically 0 [108].

The second consideration relates to aliasing in the bispectral domain. If the sampling rate is in the range \(3B_1 \leq f_s \leq 3B_2\), the bispectrum in a region beyond
the non-redundant region shown in Figure 4.2(b) is non-zero. Therefore to ensure there is no aliasing in the bispectral domain, the sampling rate must be greater than $3B_2$ where $B_2$ is the highest frequency component of the continuous time signal.

The centre frequency and bandwidth of the GPR antennas is $1.4\,\text{GHz}$ and $1.1\,\text{GHz}$ respectively. For this case, $B_1 = 850\,\text{MHz}$ and $B_2 = 1.95\,\text{GHz}$. As $B_2$ is more than $2B_1$, the bispectrum is not identically zero in the non-redundant region. Further, as the equivalent sampling rate of the GPR sampling system ($111.3\,\text{GHz}$) is greater than $3B_2$, it is assumed that the GPR signal is not aliased in the bispectral domain and hence the equations of Section 4.3 apply. Figure 4.6 shows the frequency domain magnitude spectrum of a raw GPR signal acquired from the test GPR system to confirm the bandwidth range specified.

![Figure 4.6: The magnitude spectrum of a raw GPR signal acquired from the test GPR system.](image)

Common pre-processing tools such as sum-of-squares normalisation may also be applied to the segmented data. This normalisation method extends the amplitude invariance to the magnitude parameter also. Further, the feature vector
Chapter 4. Early-Time Feature Vector for Near-Surface Processing

dimensionality can be increased by using more \( a \) values. The effects of the various pre-processing options and parameters on feature values for sub-surface scenarios are evaluated in the next section using synthetic data.

4.4.3 Class Separation

Experiment

The FDTD method presented in Section 3.4.1 was used to generate synthetic data for evaluating the proposed early-time feature vector. Two datasets were generated to investigate the discrimination between feature values for data acquired from variations in both top layer thickness and bottom layer electromagnetic parameters.

The media under investigation comprised of two layers. The upper layer represented the coal seam floor and hence the electromagnetic properties were set to values similar to those of coal. The thickness of the upper layer was increased from 1 cm to 10 cm in 0.5 cm increments. The lower layer represented the surrounding stratum to be sensed. The electromagnetic properties for this lower layer were changed systematically to analyse the effects on the feature values.

The first dataset was acquired where the contrast in electromagnetic parameters between the two layers was relative permittivity. The relative permittivity of the lower layer was increased from 6 to 24 while the conductivity for both layers was zero. The second dataset was acquired with both permittivity and conductivity contrasts however the changing parameter was conductivity. The conductivity of the lower layer was increased from 0 to 500 mS/m. Refer to table 4.1 for the electromagnetic parameters used to generate the synthetic data and Figure 4.7 for the layout of the synthetic coal seam floor. The antennas were in the surface-coupled configuration for both experiments. Figure 4.8 shows a synthetic GPR trace generated from just a single layer of coal. This represents the case of an infinitely thick coal layer. The trace comprises of clutter from antenna crosstalk and ringdown.
4.4. Proposed Early-Time Feature Vector

Table 4.1: Relative permittivity and conductivity values for the synthetic data. The first two columns contain the relative permittivity and conductivity of the top (coal) layer for both experiments. The middle two columns contain the values for the change in permittivity experiment. The last two columns contain the values for the change in conductivity experiment.

<table>
<thead>
<tr>
<th>Top Layer</th>
<th>Bot Layer - Exp 1</th>
<th>Bot Layer - Exp 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon_{r1}$</td>
<td>$\sigma_1 (mS/m)$</td>
<td>$\varepsilon_{r2}$</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>24</td>
</tr>
</tbody>
</table>

Figure 4.7: FDTD representation of synthetic coal seam floor.
Figure 4.8: FDTD scan using surface-coupled antennas with top layer only.

The raw synthetic GPR data for the first experiment is shown in Figure 4.9 where the conductivity of the bottom layer was 0. Figure 4.9(a) is for a fixed relative permittivity of both layers and a changing top layer thickness. The $y$ axis shows the top layer thickness for a given trace. The reflection from the bottom of the coal layer is clearly visible as its time delay increases proportional to the top layer thickness. This can be seen from the wavelet peak that is shifting right from sample 125 in the bottom trace to sample 250 in the top trace. The greater time delay in the wavelet peak is due to the thicker top layer.

Figure 4.9(b) shows the synthetic GPR data for a fixed top layer thickness and increasing bottom layer relative permittivity. The change in signal shape is much more subtle than the case with changing top layer thickness. As the relative permittivity of the bottom layer increases, the peak between samples 150 and 200 increases in amplitude, which is the reflection from the interface between the top and bottom layers. This change can be seen by comparing the peak amplitude of a given trace between samples 150 and 200 with the peak amplitude at sample 250 of the same trace.

The raw synthetic GPR data for the second experiment is shown in Figure 4.10 where the relative permittivity of the bottom layer was constant while the
4.4. Proposed Early-Time Feature Vector

Figure 4.9: Figure shows plots of synthetic data with lossless media for (a) changing top layer thickness with fixed permittivity and (b) changing permittivity of bottom layer with top layer thickness of 5 cm.

conductivity increased. Figure 4.10(a) is for a fixed permittivity and conductivity of both layers and a changing top layer thickness. As for the first experiment, the reflection from the bottom of the coal layer is clearly visible as its time delay increases proportional to the top layer thickness.

Figure 4.10(b) shows the synthetic GPR data for a fixed top layer thickness and increasing bottom layer conductivity. Once again, the change in signal shape is much more subtle than the case with changing layer thickness.

Figure 4.10: Figure shows plots of synthetic data with conductive bottom layer for (a) changing top layer thickness with fixed permittivity and (b) changing bottom layer conductivity with top layer thickness of 5 cm.
Permittivity Contrasting Layers

Figure 4.11 shows the feature vector values for varying both top layer thickness and bottom layer relative permittivity. In Figure 4.11(a), each line represents the feature values for a constant top layer thickness. The feature values extend outwards from the centre (labelled as $\varepsilon_r = 6$) as the relative permittivity increases to $\varepsilon_r = 24$. Figure 4.11(b) shows the same feature values however the points are joined along lines of constant bottom layer relative permittivity. As the top layer thickness increases, the feature values translate from the left side (labelled as 1 cm) in a clockwise direction around to the point where the features converge (labelled as 10 cm).

The figures indicate that the separation between feature values is greater for thinner top layers and higher relative permittivity contrasts. The reason for this is as follows. Consider the cluster labelled 8–10 cm. For this range of top layer thickness, the reflected echo from the interface between the top and bottom layers is delayed such that it has little effect on the early-time signal. Hence feature values of this layer thickness are very close and hence it would be difficult for a classifier to discriminate between the classes. As the top layer thickness decreases, the time delay of the reflected echo from the interface is less and has more effect on the early-time signal. The effect on the early-time signal will continue to increase as the time delay decreases which is why the feature values are more sparse for a thinner top layer thickness such as 1–3 cm.

The effects of four processing variations in the segmentation and feature extraction steps have been investigated. These steps are window type, window length, normalisation and number of $a$ parameters.

The window types considered in this investigation were the rectangular and Hamming windows. The choice of window for a given task is predominantly determined by the stability of the end sections of the early-time GPR signal. This can be determined either emperically or from a plot that shows the variance of each sample. A noise analysis of the test GPR system is presented in Appendix B. Figure B.2 shows the variance of free space and coal surface scans. The variance between samples 100-200 is very small compared to the peaks of the variance plots which are at approximately sample 30. If the data was segmented from the
first sample to somewhere between samples 100 and 200, the rectangular window has superior class discrimination as the endpoints of the segmented signal are stable. If however the end points were to be at a highly varying sample such as sample 30, then the Hamming window provides more stability in the clusters yet less discrimination. More stability is provided because the variations are attenuated as the amplitude of the tails of the Hamming window taper down towards zero compared with the amplitude of the window peak. There will be less discrimination for a Hamming window with a given window length compared with a rectangular window of the same length as less information is being retained due to, similarly, the lower amplitude of the Hamming window tails. This can be seen by comparing Figures 4.11(a) and 4.12(a) which show the feature values for the rectangular and Hamming windowed data. The early-time signal windows using the rectangular window has better feature separation than the data processed with the Hamming window. From a practical perspective, however, the final choice of window type can also be made empirically from inspection of the training data examples.

The objective of the normalisation step is to extend amplitude invariance to the integrated bispectral magnitude parameter. The effect of sum-of-squares normalisation can be seen by comparing the sub-figures of Figure 4.12. Figure 4.12(a) is for the Hamming window without normalisation and 4.12(b) for the same window with normalisation. The window length in both cases is 160 samples. The phase parameter is unchanged however the magnitude has changed. Therefore sum-of-squares normalisation has no effect on the integrated phase, however the integrated bispectral magnitude is affected. One limitation however, is that for this case the feature path of the normalised features overlap which will degrade the performance of the classifier.

The number of \( a \) parameters affects the dimensionality of the feature vector and classifier. More dimensions will increase the computational complexity and can also cause the pattern recognition to be tuned to a given dataset as opposed to being more generalised. Therefore it is common practice to select a feature space dimension that provides satisfactory results for a given application. For the integrated bispectral parameters, it is usual to select \( a = 1 \) when a single feature
vector is required. It is however necessary to consider the cluster discriminability for values of $a < 1$. Figure 4.13 shows the feature values for $a = 0.25, 0.5, 0.75$ and 1 for comparison. For this experiment, the feature values obtained with $a < 1$ are closer than or similar to the $a = 1$ case. Therefore a performance increase is not expected if the feature dimensionality is increased.

The window length has a direct relationship with range sensitivity and resolution. Figure 4.14 shows the Euclidean distances for varying window lengths. The $x$ axis represents the top layer thickness and the $y$ axis represents the bottom layer relative permittivity. The image intensity scale represents the Euclidean distance between the feature value of the given point and the 10 cm layer thickness with $\varepsilon_r = 6$. The features were computed using a rectangular window without normalisation. Figure 4.14(a) shows the distances from a window length of 100 samples. This part of the early-time signal is only affected when the top layer is thin. The black region of this figure indicates that the Euclidean distance between the feature values for a thick top layer (10 cm) and moderately thick top layer (6–10 cm) is small for all bottom layer relative permittivities. The light area indicates that there is good feature separation from the thick top layer case. As the window length increases, more energy reflected from the bottom layer is included for processing. Hence the brighter area spreads towards the increasing top layer thickness direction indicating that there is more feature separation for a thicker top layer. These figures also show that the Euclidean distance increases as the bottom layer relative permittivity increases.
4.4. Proposed Early-Time Feature Vector

Figure 4.11: Feature values for rectangular window with length 160 samples showing paths of constant (a) top layer thickness and (b) bottom layer relative permittivity. The data is the same for both plots.
Figure 4.12: Feature values for Hamming window with length 160 samples with and without sum-of-squares normalisation.
Figure 4.13: Feature values for rectangular window with 160 samples indicating effect due to varying $a$ parameter.
Figure 4.14: Intensity representation of Euclidean distances between a given point and the top layer thickness of 10cm point as the window length increases. Features computed with a rectangular window for varying bottom layer relative permittivity.
Conductivity Contrasting Layers

Figure 4.15 shows the feature vector values for varying both top layer thickness and bottom layer conductivity. As for the permittivity contrasting experiment, Figure 4.15(a) shows lines representing the feature values for a constant top layer thickness. The feature values extend from the bottom left corner (labelled as $\sigma = 0$) upwards and toward the right as the bottom layer conductivity increases to 500 mS/m as labelled. In Figure 4.15(b), each line represents the feature values for a constant bottom layer conductivity as indicated. The feature values traverse along the lines extending in a clockwise direction from the left side as the top layer thickness increases. The feature values converge at the point labelled 7–10 cm, which represents a thicker top layer. As for the contrasting permittivity experiment, the figures indicate that the separation between feature values is greater for thinner top layers and higher conductivity contrasts. The cause of this is the same as explained for the contrasting permittivity experiment.

The feature values for the rectangular and Hamming windows are shown in Figures 4.15(a) and 4.16(a). For the change in conductivity experiment, the rectangular window features have greater class separation over those with the Hamming window. This can be seen by comparing the figures and taking into account the scale difference on the magnitude axis.

As for the permittivity contrasting experiment, Figures 4.16, 4.17 and 4.18 show the effect on the feature values due to normalisation, varying $a$ parameters and changing window length respectively. The discussion for these is the same as for the changing permittivity experiment and hence is not repeated here.
Figure 4.15: Feature values for rectangular window with length 160 samples showing paths of constant (a) top layer thickness and (b) bottom layer conductivity. The data is the same for both plots.
4.4. Proposed Early-Time Feature Vector

Figure 4.16: Feature values for Hamming window with length 160 samples with and without sum-of-squares normalisation.
Figure 4.17: Feature values for rectangular window with 160 samples indicating effect due to varying $a$ parameter.
4.4. Proposed Early-Time Feature Vector

Figure 4.18: Intensity representation of Euclidean distances between a given point and the top layer thickness of 10cm point as the window length increases. Features are computed with a rectangular window for varying bottom layer conductivity.
4.4.4 Noise Reduction

The most common method of reducing noise power in GPR processing is ensemble averaging of the raw GPR traces. An alternative to ensemble averaging of the raw data is to average in the feature domain. Feature averaging involves computing the mean of the feature vector for a given dataset.

Table 4.2 shows the bias and variance of the magnitude and phase features for the two noise reduction approaches - ensemble averaging and feature averaging. \( N \) represents the number of realisations (ensemble averaging) or feature values (feature averaging) used so that a direct comparison between the approaches can be made. The features were computed from a synthetic GPR scan with additive noise. The noise process was measured from the test GPR system and is shown in Figure B.1. The noise statistics are presented in Appendix B. The features were standardised to unit variance for the \( N = 1 \) case.

The table indicates that the variance decreases faster with feature averaging than ensemble averaging. The bias however decreases for ensemble averaging but remains constant for feature averaging. Therefore it is concluded that feature averaging will tend to generate biased estimates of the feature values being estimated, however the spread of the clusters will decrease. In a practical sense, this will result in a systematic error in the final metric being estimated.

4.5 Summary

This chapter has described a novel feature vector that characterises the early-time signal of 1D GPR data. A full evaluation of the feature vector and various processing options has been presented using synthetic GPR data. Two noise reduction methods have been compared for the approach indicating that for the noise process of the test GPR system, feature averaging reduces noise variance more than ensemble averaging. It is also shown that any bias in the features remains after feature averaging, however this decreases with ensemble averaging. The next chapter presents a novel approach to estimate the thickness of a thin top layer using GPR and the feature vector proposed in this chapter.
Table 4.2: Bias and variance statistics of the magnitude and phase features for ensemble averaging and feature averaging. Features were computed from a synthetic GPR signal with additive noise measured from the test GPR system.
Chapter 5

Thin Layer Thickness Estimation

5.1 Introduction

This chapter proposes a method to estimate the thickness of a thin top layer using the feature vector developed in Chapter 4 with a nearest neighbour classifier. The thin layer thickness estimator was evaluated under two imaging scenarios with both synthetic and real data. The top layer was coal for both experiments. The bottom layer was shale for the first experiment and clay for the second. The antennas were surface-coupled for both experiments.

To begin the chapter, a review of a radar signal model and the traditional processing methods for layer thickness estimation with GPR is presented. The proposed thin layer thickness estimator is then presented which includes specific parameters selected for the early-time feature vector and the classifier. This is followed by the results of the thickness estimator and discussion.

5.2 Layer Thickness Estimation

5.2.1 Radar Signal Model

For the case of horizontal-layered media, the received signal of an impulse GPR system can be modelled as the superposition of attenuated and delayed replicas of a known signal $s(t)$ for each interface plus clutter components such as the background signal and noise [101]. Figure 5.1 shows a graphical representation of
Chapter 5. Thin Layer Thickness Estimation

this model for both air-coupled and ground-coupled antenna configurations. The mathematical form of this model is

\[ r(t) = \sum_{m=0}^{M} a_m s(t - \tau_m) + b(t) + n(t) \]  

(5.1)

where \( r(t) \) is the received signal, \( s(t) \) is the transmitted signal, \( M \) is the total number of sub-surface interfaces, \( a_m \) is the peak amplitude of the reflection from the \( m^{th} \) sub-surface interface, \( \tau_m \) is the time delay of the reflection from the \( m^{th} \) sub-surface interface, \( b(t) \) is the background signal and \( n(t) \) is additive noise. The additive noise is assumed to be Gaussian.

The model in (5.1) is for air-coupled antennas. This model can be extended to ground-coupled antennas by setting the term for the \( 0^{th} \) sub-surface interface to zero, i.e. \( a_0 \) is zero. The reflection from the \( 0^{th} \) sub-surface interface is the ground-bounce reflection from the air-ground interface and does not apply for ground-coupled antennas.

The amplitude of the reflected signal from the \( m^{th} \) interface, \( a_m \), is a function of the interface reflection coefficient, which in turn, is a function of the conductivity and permittivity contrasts of the layers. This is based on the assumption that the media are non-magnetic. For the three layers shown in Figure 5.1, the conductivities are represented as \( \sigma_1, \sigma_2, \sigma_3 \) and permittivities are \( \varepsilon_1, \varepsilon_2 \) and \( \varepsilon_3 \) respectively. The permittivity for free-space is shown as \( \varepsilon_0 \) and is \( 8.854 \times 10^{-12} \) [92]. Information regarding the estimation of \( \varepsilon \) from signal amplitudes, \( a_m \), is given in Section 5.2.3.

In many situations the model is simplified by assuming the effect of conductivity is negligible. Relaxing this assumption complicates any physical GPR model as it has been shown that received echoes from boundary changes in conductivity can be interpreted as boundary changes in permittivity [56]. Additionally, the amplitude is also affected by propagation phenomena such as antenna gain and geometric spreading.

This signal model relies on knowledge of the signal transmitted by the GPR system. As it is often impractical to measure the transmitted signal, it is common practice to estimate \( s(t) \) using the reflection from a metal plate during calibration. This estimate of the transmitted signal is used by a matched filter to detect the
target echoes.

The time delay of each detected target echo, $\tau_m$, is estimated from the peak times of the matched filter output or specific points on a detected pulse. This is described in Section 5.2.4. The layer thicknesses are then estimated using the time delay estimates and propagation velocities (either assumed or estimated). These stages along with the background signal, $b(t)$, are discussed in the following sections.
5.2.2 Echo Detection: The Matched Filter

The matched filter is the basis for design of almost all radar receivers as it is the optimal method of detecting a deterministic signal in additive Gaussian noise. The frequency response of the matched filter is the complex conjugate of the transmitted signal. The impulse response of the matched filter is a time reversed and shifted version of the transmitted signal \([100]\).

In discrete-time systems, the matched filter can be implemented as the replica correlator. The replica correlator correlates the received waveform with a reversed time shifted version of the transmitted signal to obtain the likelihood ratio function. For discrete-time sequence \(x(n)\), the likelihood ratio function of the replica correlator is computed using \([49]\)

\[
L(n) = \exp\left\{ -\frac{1}{2\sigma^2} \left( \sum_{m=0}^{M-1} s^2(m) - 2 \sum_{m=0}^{M-1} x(n - m) s(m) \right) \right\} \tag{5.2}
\]

where \(n\) is the sample number, \(s(m)\) is the estimate of the transmitted signal with length \(M\), \(x(n)\) is the received signal and \(\sigma^2\) is the noise variance (average noise power). A target is considered as present if the likelihood ratio is greater than the detector threshold. If the threshold is set too low, false alarms from clutter will result. If the threshold is set to high, false alarms due to noise and clutter will not occur however the chance of detecting a weak target will decrease. Figure 5.2 shows a typical output of a correlation matched filter. Echoes A, B and C are from targets whereas D is due to clutter. The top dashed line represents the threshold that governs whether an echo is detected. For the given threshold, echoes A and B would be successfully detected and echo C (weak target) would be missed. If the threshold was lowered such that C was detected, a false alarm will result as D would also be detected.

An important parameter for layer thickness estimation is the time delay of the detected echo. The time delay of the \(m^{th}\) detected echo, \(\tau_m\), is estimated from the peak time of the matched filter output.
5.2. Layer Thickness Estimation

Figure 5.2: Image shows the output of a correlation matched filter as a function of time. Echoes A, B and C are from targets whereas D is due to clutter.

5.2.3 Propagation Velocity Estimation

The propagation velocity is a function of the relative permittivity of the medium and is

\[ v = \frac{c}{\sqrt{\varepsilon_r}} \]  \hspace{1cm} (5.3)

where \( c \) is the electromagnetic wave propagation velocity in free space (\( 3 \times 10^8 \) m/s) and \( \varepsilon_r \) is the dielectric constant of the medium. For a lossless and relatively smooth surface using air-coupled antennas, the dielectric constant of the top layer can be estimated from the amplitude of the ground-bounce reflection, \( a_0 \), and a metal plate reflection, \( a_{mp} \),

\[ \varepsilon_1 = \left[ \frac{1 + a_0/a_{mp}}{1 - a_0/a_{mp}} \right]^2. \]  \hspace{1cm} (5.4)

The dielectric constant of deeper layers such as \( \varepsilon_2 \) and \( \varepsilon_3 \) in Figure 5.1 may be estimated using a technique known as layer-stripping inversion and equations as described in [101].

As the transmitted electromagnetic energy is subject to spreading as the wave propagates, the metal plate reflection used for propagation velocity estimation should be acquired with the antennas at the same height as for the survey. This involves either constraining the operating range of the antenna height during acquisition and/or acquiring the metal plate calibration scans at a wide range of antenna heights.

Techniques exist to estimate the wave propagation velocity in the top layer using ground-coupled antennas. Two examples of these are common mid-point and wide-angle reflection and refraction [32]. These methods require the repeated separation of the antennas in a controlled fashion.
In the absence of measured dielectric constant information, it is common practice to assume a propagation velocity based on published values of the dielectric constant for the target media.

### 5.2.4 Thickness Estimation

The thickness of the first layer using ground-coupled antennas is estimated using

\[
d = \sqrt{\left[ (t_1 - t_d + t_c) \frac{v}{2} \right]^2 - \left( \frac{x}{2} \right)^2}
\]

where \(d\) is the layer thickness estimate, \(t_1\) is the positive peak time of the first interface reflection, \(t_d\) is the negative peak time of the direct signal, \(t_c\) is a time correction factor, \(v\) is the wave propagation velocity through the first layer and \(x\) is the transmitter to receiver antenna separation. The geometry of this model is shown in Figure 5.3.

![Figure 5.3: Geometry of layer thickness estimation signal model for ground-coupled antennas.](image)

For air-coupled antennas, the first step is to estimate the height of the antennas. The antenna height is estimated using

\[
h = \sqrt{\left[ (t_1 - t_d + t_c) \frac{c}{2} \right]^2 - \left( \frac{x}{2} \right)^2}
\]

where \(h\) is the antenna height and \(c\) is the propagation velocity for free-space. The time correction factor, \(t_c\), can be estimated by repeatedly acquiring data with air-coupled antennas at different but known antenna heights, \(h\), above the ground surface. The value for \(t_c\) that minimises the squared error of the estimated
5.2. Layer Thickness Estimation

The time correction value obtained using this procedure for the GPR system described in Chapter 3 is 288.4 ps.

A model that represents the return propagation time to the first sub-surface interface with air-coupled antennas is [59]

\[ t_{\text{tot}} = \left( \frac{2}{c} \right) \sqrt{\left( \frac{x}{2} - a \right)^2 + h^2} + \left( \frac{2}{v} \right) \sqrt{a^2 + d^2} \]  

(5.7)

where \( t_{\text{tot}} = t_1 - t_d + t_c \), \( a \) is the distance between the midpoint of the antennas and surface location where the incident ray refracts through the surface and \( d \) is the top layer thickness. Both parameters \( a \) and \( d \) are unknown and need to be solved independently. Using the wave propagation velocity estimate through the top layer (5.3) and angle of refraction, a fourth order polynomial in \( a \) can be solved numerically [59]. By rearranging (5.7), the top layer thickness can be estimated with air-coupled antennas using

\[ d = \sqrt{\frac{v^2}{4} \left\{ t_{\text{tot}} - \frac{2}{c} \sqrt{\left( \frac{x}{2} - a \right)^2 + h^2} \right\}^2 - a^2} \]  

(5.8)

If the antenna module is either monostatic or the antenna separation is small such that \( x \) and subsequently \( a \) are both zero, (5.8) can be simplified to

\[ d = \frac{v}{2} \left( t_{\text{tot}} - \frac{2h}{c} \right) \]  

(5.9)

The geometry of the air-coupled model is shown in Figure 5.4.

5.2.5 Clutter and Background Subtraction

In the context of GPR, clutter can be defined as “those signals that are unrelated to the target scattering characteristics but occur in the same sample time window and have similar spectral characteristics to the target wavelet” [30]. For the case of layer thickness estimation, clutter consists of the direct transmitter to receiver signal (also called breakthrough or crosstalk) and antenna ringdown. The ringdown is caused primarily by the re-radiated fields due to currents reflecting within the antenna and associated structures [5]. The crosstalk and antenna ringdown are lumped as the background component \( b(t) \) in (5.1). These components swamp
Chapter 5. Thin Layer Thickness Estimation

Figure 5.4: Geometry of layer thickness estimation signal model for air-coupled antennas.

The early part of the received signal and can be seen as the horizontal banding in Figures 5.5(b) and 5.6(a).

The most common approach used to minimise the effects of the background signal is to subtract the mean trace. This is called background subtraction and can be implemented as either a moving average filter over a defined spatial window or the mean of an entire B-scan. This pre-processing tool can give satisfactory results when imaging point reflectors, however the response from an interface may be filtered out using this process. Other attempts to reduce clutter include subtracting a free-space calibration scan which incorporates the crosstalk and antenna ringing. This method is often unsuccessful as the antenna centre frequency decreases when the antennas are in contact with the ground which broadens the transmitted pulse. Residual error usually results after the process of background subtraction using the mean trace or free-space calibration. This residual error can be falsely detected by the matched filter as a target which increases the false alarm rate. In many cases these background removal techniques yield favourable results when one is primarily interested in viewing the data. Figures 5.5 and 5.6 show the enhancement in B-scan data using mean trace background subtraction for both point and plane reflector cases. However if quantitative information
such as layer thickness is required, it is common practice to treat the early part of
the signal as a region where reflections are undetectable and impose a minimum
thickness operating range.

![Figure 5.5: Raw and background removed B-scan of point reflector.](image)

(a) Raw data of point reflector  (b) Background removed data of point reflector

![Figure 5.6: Raw and background removed B-scan of plane reflector.](image)

(a) Raw data of plane reflector  (b) Background removed data of plane reflector

Figure 5.5: Raw and background removed B-scan of point reflector.

Figure 5.6: Raw and background removed B-scan of plane reflector.
5.3 Proposed Thin Layer Thickness Estimator

5.3.1 Segmentation and Feature Extraction

Chapter 4 showed that the proposed early-time feature vector contains valuable information for discriminating between different cases of sub-surface conditions. The cluster plots shown in Section 4.4.3 may be used as a guide to select the feature vector parameters for a set task. As the objective of the thin layer thickness estimator is to discriminate between different classes of layer thickness, other information such as the relative permittivity or conductivity of the bottom layer is not important. Therefore feature vector parameters should be chosen such that they provide discrimination between layer thickness classes.

Figures 4.11(a) and 4.12(a) show clear discrimination between layer thickness the case of up to approximately 4–5 cm with synthetic data. The feature values in these figures were computed from a window length of 160 samples without sum-of-squares normalisation, $a = 1$ and for both Hamming and rectangular windows. The features from the Hamming window are closer together in the range above 3 cm. The sum-of-squares normalisation step modifies the magnitude parameter such that the classes are no longer unique. In this case, the result of incorporating the Hamming window and/or normalisation reduce the discriminability of the classifier. Therefore the parameter set deemed most suitable for the thin layer thickness estimation task using the given synthetic data is the rectangular window of 160 samples without normalisation.

5.3.2 Classifier

The nearest neighbour method is an elementary but popular classification tool [35]. The algorithm is based on classifying a given set of test data into the closest training class in the feature space. The most common distance metric for determining the class is the Euclidean distance.

The training phase involves mapping training data into a multi-dimensional feature space where the number of dimensions is equal to the number of elements in the feature vector. The feature space is partitioned into different classes where
each class has a label corresponding to the known training data. In the testing phase, the test data is transformed into points in the feature space. Each test point is allocated the label of the training point that is the closest based on the Euclidean distance. The nearest neighbour method can naturally be extended such that a given test point is assigned to the class that is the most frequent of the \( k \) nearest neighbouring test points. This is called the \( k \)-Nearest Neighbour (\( k \)-NN) algorithm [35]. The nearest neighbour algorithm is a special case of the \( k \)-NN method where \( k = 1 \).

5.4 Experiment 1: Coal-Shale Interface

A common overburden material for an underground coal mine is shale. It is therefore a suitable media for testing the proposed thin layer thickness estimator. A graphical representation of this imaging scenario is shown in Figure 5.7.

![Figure 5.7: Image showing the experimental setup for the coal-shale interface evaluation.](image)

5.4.1 Synthetic Data

Synthetic data was generated using the FDTD simulator discussed in Section 3.4.1. The model was excited with the source signal presented in Section 3.4.2 so that the effects of clutter such as antenna ringing and crosstalk were included. The sub-surface scenario investigated was the two layer model as per the experiment
described in Section 4.4.3 for evaluating the early-time feature vector. The top layer was coal and the bottom layer was shale. The thickness of the top layer was systematically increased from 0.5 cm to 10 cm in 0.5 cm increments. The dielectric properties of the coal and shale used in the experiment were measured from the testbed material. The dispersive nature of the testbed media were modelled using the Debye relaxation model. The model parameters used to model the dispersion are shown in Table 3.2.

**Processing**

The initial processing steps included DC offset removal and segmentation using a rectangular window with 160 samples. The features were standardised [35] to ensure equal weighting between individual feature vector components.

The feature vectors were extracted from the clean synthetic data with known top layer thickness. As these feature values were obtained from data with 0.5 cm thickness increment, a cubic B-spline [82] was fitted to interpolate locally between the feature vectors obtained from the clean data in increasing layer thickness. The points on the spline from the interpolation were merged with the feature vectors from the synthetic data with known layer thickness to form the full training set. This allows the transition of the thickness estimate towards a continuous variable with millimetre resolution rather than discrete with half centimetre spacing. The thickness estimate is obtained from the single class along the spline contour closest to the test feature vector, which is nearest neighbour classification. Figure 5.8 shows the training feature vectors with known layer thickness at 0.5 cm increments, the cubic spline and the interpolated points. The point for the 0.5 cm case is shown in the bottom left corner. The path of increasing top layer thickness progresses in the clockwise direction towards the 10 cm point as per the labels.
5.4. Experiment 1: Coal-Shale Interface

Figure 5.8: Plot showing variation of feature vector for synthetic coal-shale interface over increasing coal layer thickness (as labelled).
Additive Noise

The effect on the thin layer thickness estimator due to the measured noise from a real GPR system was investigated. The measured noise statistics are described in Appendix B. Figure 5.9 shows the feature values estimated from the noisy synthetic data. The thickness estimates from regions of high concentration of training points (either measured or interpolated) are subject to higher variance. Inspection of Figure 5.8 reveals that thickness estimates between 3 and 4 cm and greater than 7 cm are subject to this phenomenon.

Figure 5.9: Plot showing feature vector clusters of noisy synthetic data acquired from the coal-shale media.
5.4. Experiment 1: Coal-Shale Interface

Results

The top layer thickness was estimated from the synthetic data with additive noise. The mean thickness estimate for each ensemble along with the 99% confidence interval is shown in Figure 5.10. As expected, the estimator variance is higher in the regions of high training feature vector concentration.

![Figure 5.10: Plot showing true layer thickness vs estimated layer thickness for synthetic coal-shale interface over increasing coal layer thickness. The errorbars represent the 99% confidence interval.](image)

As described in Section 4.4.4, noise power may be decreased through averaging. This in turn reduces the variance of the thickness estimates. Effective comparative measures of estimator performance with respect to noise are the bias, variance and mean squared error (MSE). These measures are shown in Figures 5.11, 5.12 and 5.13 for the true thickness values of 2 and 4 cm as the signal to noise ratio (SNR) is increased via ensemble averaging. The figures indicate that in terms of the bias, variance and MSE, the estimator is superior when the training feature values are less concentrated as for the 2 cm case.
Chapter 5. Thin Layer Thickness Estimation

Figure 5.11: Plot showing absolute value of bias of estimator for top layer thicknesses of 2 cm and 4 cm.

Figure 5.12: Plot showing variance of estimator for top layer thicknesses of 2 cm and 4 cm.
Figure 5.13: Plot showing MSE of estimator for top layer thicknesses of 2 cm and 4 cm.
Chapter 5. Thin Layer Thickness Estimation

5.4.2 Real Data

The previous section presented an evaluation of the proposed thin layer thickness estimator using synthetic GPR data. In this section the estimator is evaluated with real GPR data acquired from the experimental testbed with a coal top layer and shale bottom layer. The testbed region under investigation in this experiment is the traverse distance from 1000 to 2400 mm in Figure 3.5.

Processing

The processing steps and feature extraction parameters for the real data experiment were identical to those for the synthetic data, i.e.: DC offset removal, rectangular window length of 160 samples and feature standardisation.

The method proposed in the previous section of interpolating between coarsely measured feature values was unsuccessful as the spline overlapped and was non-unique. This resulted in ambiguity in some of the thickness estimates. Therefore the classifier was trained using feature vectors from some of the measured data with known top layer thickness. The real GPR data used for training was ensemble averaged which resulted in one trace for each class. The feature values used to train the classifier for the coal-shale media are shown in Figure 5.14.

The real GPR data used for testing the classifier was acquired at positions near the measured points but in the direction normal to the measured line. This was so that it can be assumed the true layer thickness is similar to the measured points but the antennas were not at the same location. Further, the test data was acquired on a different day to the training data.

To enable a comparison between the proposed thin layer thickness estimator and the traditional thickness estimator for ground-coupled antennas based on the signal model detailed in Section 5.2.1, the test data was also processed using the traditional processing strategies. The traditional approaches predominantly rely on the subtraction of a background calibration signal and matched filtering to detect the echo from the sub-surface interface. The processing stages for this approach include the subtraction of a calibration signal, time varying gain to enhance the contrast of deeper reflections, and matched filtering to detect the target echo. The relative permittivity used in calculations for time delay to layer
5.4. Experiment 1: Coal-Shale Interface

Figure 5.14: Plot showing feature vectors for real data with coal-shale interface used to train classifier. The labels denote the measured top layer thickness.

thickness conversion was 4.4 as per Table 3.1.

Results

Figure 5.15 shows the measured and estimated top layer thickness for both the proposed estimator and the matched filter approach. The error bars represent the 99% confidence interval of the estimates. The results of the proposed thin layer thickness estimator were reasonably accurate when the layer is thin however the error is substantial when the layer is not thin. Conversely, the matched filter is unable to accurately estimate the layer thickness when the layer is thin. However it is reasonably successful for the thick layer case. In the case shown, the crossover point between the thin and thick layer is approximately 5 cm.

The effect of noise reduction on both thickness estimators was investigated. The main interest was to compare the estimators in their corresponding valid operating ranges. The bias, variance and MSE were measured for three different noise reduction approaches. The first was reducing noise on the matched filter
Chapter 5. Thin Layer Thickness Estimation

thickness estimator using ensemble averaging. The second and third approaches were using ensemble averaging and feature averaging on the thin layer thickness estimator. These two approaches are discussed in Section 4.4.4.

Figures 5.16 through to 5.18 shows SNR as the independent variable. The SNR for the test GPR system with no averaging is approximately 27 dB. This is shown as the lowest SNR value in these figures.

Figure 5.16 shows the bias as a function of SNR for both estimators. In Figure 5.16(a), the top layer thickness is 2.5 cm, hence the thin layer thickness estimator is operating in its valid operating range. Firstly, this figure indicates that the bias is less for the proposed estimator. Secondly, this figure also indicates that the bias for the matched filter thickness estimator is significant compared with the actual measured layer thickness. Conversely, Figure 5.16(b) indicates that when the top layer is thicker, the bias of the matched filter thickness estimator is significantly less than the thin layer thickness estimator. Finally, both of these figures indicate that there is no significant improvement in the bias for both estimators as the SNR increases, irrespective of the averaging technique, actual measured layer thickness and whether the layer is thick or thin.

Figure 5.17 shows the variance as a function of SNR for both estimators. Figure 5.17(a) shows that for a measured thickness of 2.5 cm, the variance for the thin layer thickness estimator is improved with averaging, however both averaging approaches generate similar results. Figure 5.17(b) indicates that when operating in its valid thickness range, the variance of the matched filter thickness estimator improves as SNR is increased. There are two other interesting points about these figures. Firstly, the variance for the both thickness estimators is very small or zero when not operating in their corresponding valid operating range. This results in the floor effect for the matched filter data in Figure 5.17(a) and for the ensemble and feature averaging of the thin layer thickness estimator in Figure 5.17(b). Secondly, these figures indicate that there is a lower limit to how much the variance can improve. The SNR at this point is approximately 38 dB for these cases.

The MSE represents a combination of both bias and variance statistics and is commonly used as a general measure of error. The MSE is shown in Figure 5.18
for both estimators. Figure 5.18(a) indicates that for a top layer of 2.5 cm, the thin layer thickness estimator outperforms the matched filter thickness estimator. Figure 5.18(b) indicates the converse that for a thick top layer such as 16 cm, the matched filter thickness estimator is superior. These results are as expected as each thickness estimator is superior when operating in its corresponding valid operating range.

Figure 5.19 shows the MSE versus the true layer thickness for the proposed thickness estimator and the matched filter based thickness estimator with no averaging. The figure reinforces the previous results that the proposed bispectral feature based thickness estimator is superior than the matched filter based approach when the layer is thin and the converse when the layer is thick. The crossover point between the two lines indicates the valid operating range boundary for each estimator.
Figure 5.15: Plots showing measured vs estimated layer thickness for real coal-shale interface with increasing coal layer thickness for the proposed and matched filter based estimators. The error bars represent the 99% confidence interval.
5.4. Experiment 1: Coal-Shale Interface

Figure 5.16: Plots showing the bias for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging.

(a) Thickness is 2.5cm

(b) Thickness is 16cm
Figure 5.17: Plots showing the variance for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging.
Figure 5.18: Plots showing the MSE for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging.
Figure 5.19: Plot showing the MSE versus true layer thickness for the proposed thin layer thickness estimator and the matched filter approach for no averaging.
5.5 Experiment 2: Coal-Clay Interface

Another geological scenario that can occur in an underground coal mine is the coal-clay case. Therefore it is worthy to investigate the performance of the thin layer thickness estimator is under these conditions. A graphical representation of the coal-clay imaging scenario is shown in Figure 5.20.

![Image showing the experimental setup for the coal-clay interface evaluation.](image)

5.5.1 Synthetic Data

Similarly to the first experiment, synthetic data was generated using the FDTD simulator. The same parameters and experimental setup was used however the bottom layer was clay.

The processing steps for this experiment were the same as for the coal-shale experiment using synthetic data. The random noise process was also used to investigate the effect of feature values due to noise. Similarly as described in Section 5.4.1, a cubic B-spline was fitted to locally interpolate between the feature values from the measured layer thicknesses. The measured points, spline and interpolated points are shown in Figure 5.21.

The trained classifiers for the synthetic coal-shale and coal-clay experiments resemble a similar path. They are however different and hence one cannot be used for both experiments with varying bottom layer materials. This restriction
is acceptable for the mining application as it is unlikely that the overburden for a given mining operation will change significantly. Further, the overburden material is usually determined during mine planning stages. Figure 5.22 shows the feature vectors of the measured layer thickness points of both experiments before standardisation for comparison.
Figure 5.22: Feature vectors overlayed for synthetic coal-shale and coal-clay interfaces with various layer thickness.
Results

The layer thickness was estimated for the coal-clay experiment. The spline for the classifier used for the coal-clay experiment is more rounded in the 4 cm region. This results in a decreased training point concentration compared with the classifier for the coal-shale experiment. Hence the variance at the 3–4 cm region is less than for the first experiment. Figure 5.23 shows the true layer thickness and estimated layer thickness for the synthetic coal-clay experiment.

Figure 5.23: Plot showing true layer thickness vs estimated layer thickness for synthetic coal-clay interface over increasing coal layer thickness. The error bars represent the 99% confidence interval.
5.5.2 Real Data

As in the coal-shale experiment, the testbed was used to acquire real GPR data from a coal-clay region for evaluation of the proposed thin layer thickness estimator. The testbed region under investigation in this experiment is the traverse distance from 0 to 1000 mm in Figure 3.5.

As in the experiment for the coal-shale scenario with real data, the data was segmented with a 160 samples rectangular window. The spline to interpolate between the measured training points was not used for the same reasons as explained in Section 5.4.2, that the spline overlapped and was non-unique. The matched filter was also used for a comparative measure.

![Feature vectors generated from real data with coal-clay interface. These values were used to train the classifier.](image)

Figure 5.24: Plot showing feature vectors generated from real data with coal-clay interface. These values were used to train the classifier.

Results

The measured and estimated top layer thickness for both the proposed estimator and the matched filter approach are shown in Figure 5.25. The results of both
the thin layer thickness estimator and matched filter thickness estimator in this
eperiment were similar to the coal-shale experiment. The thin layer thickness
estimator was reasonably accurate when the layer is thin however the error is
substantial when the layer is not thin. Conversely, the matched filter is unable to
accurately estimate the layer thickness when the layer is thin, but is reasonable
when the layer is thick.

The effect of noise reduction on both thickness estimators was also investigated
for the coal-clay experiment. The bias for both estimators is shown in Figure 5.26.
The variance is shown in Figure 5.27 and the MSE is shown in Figure 5.28.

The bias, variance and MSE results for the coal-clay experiment are similar to
that for the coal-shale experiment. The comments from the coal-shale experiment
also apply here. However the point where increasing the SNR doesn’t continue
to decrease the variance is approximately 34 dB for the coal-clay experiment.
Figure 5.25: Plots showing measured vs estimated layer thickness for real coal-clay interface with increasing coal layer thickness for the proposed and matched filter based estimators. The error bars represent the 99% confidence interval.
Figure 5.26: Plots showing the bias for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging.
5.5. Experiment 2: Coal-Clay Interface

Figure 5.27: Plots showing the variance for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging.
Figure 5.28: Plots showing the MSE for the proposed thin layer thickness estimator and the matched filter approach as the SNR is increased through averaging.
Figure 5.29: Plot showing the MSE versus true layer thickness for the proposed thin layer thickness estimator and the matched filter approach for no averaging.
5.6 Discussion

The cubic splines used for processing the synthetic data have areas of high concentration of feature values. These regions, between 2.5–3 cm and greater than 6 cm for the coal-shale experiment, are susceptible to significant estimation error. Greater layer thicknesses result in the target reflection being delayed along the GPR trace. Hence the early time segment of the trace which is extracted for processing has less energy from the target reflection when the layer is thick. As the layer thickness increases, the feature vector will converge around the points labelled 10 cm (ref Figure 5.8). For the proposed processing technique, this approaches the case when the lower layer is out of range of the GPR system. Hence the current technique is unreliable when the top layer is thick and traditional techniques such as the matched filter may be the processing tool of choice.

Even though the splines for the classifier of the proposed estimator were omitted for the real data experiments, the results were still superior than the matched filter approach when the layer was thin. The downfall of the matched filter is due to the reflection from the interface between the top and bottom layer not being successfully detected when the layer is thin. Rather residual error from the background subtraction process is being detected as a target echo. This results the layer thickness being over-estimated in this case.

The bias, variance and MSE figures indicate a valuable performance measure of the estimators for both thin and thick layers. These statistics are shown in Figures 5.16, 5.17 and 5.18 for the coal-shale experiment and 5.26, 5.27 and 5.28 for the coal-clay experiment. These figures indicate the following points.

Firstly, that the bias of the thickness estimator that is operating in its respective valid operating range is less than the the estimator not operating in its valid operating range.

Secondly, the bias doesn’t decrease as the SNR is increased irrespective of averaging technique, actual measured layer thickness and whether the layer is thick or thin. It is, therefore, more likely that any error in the thickness estimates is due bias.

Thirdly, the variance of the thickness estimator decreases with averaging when
operating in its respective valid operating range. However the variance is either small or zero when not operating in its valid operating range and hence does not improve with averaging.

Taking these comments into account, it is concluded that both the proposed thin layer thickness estimator and matched filter thickness estimator are biased estimators. The bias may be attributed to several factors such system, environment or measurement. System dependant bias could be due to non-linearities in the down conversion and sampling process which converts the highly sampled 111.3 GHz GPR signal down to 28.8 kHz. Measurement bias could be due to slight inconsistencies during the layer thickness measurement process. Environment bias could be due to the variability and inhomogeneties in the media. Further, reflections may be scattered and received from neighbouring regions or sidewalls in the testbed that affect the waveform. From a practical perspective, it is expected that the contour of the thin sub-surface layer will not change as dramatically as in the testbed however this might not always be the case. Therefore environmentally induced bias should be limited in practice.

It is also concluded that a given estimator is only effective when the top layer thickness is within the valid operating range. For the GPR system and testbed facility used in the testing of the estimators, the operating range of the thin layer thickness estimator is from approximately 1 cm to between 4 and 5 cm. Further, the operating range of the matched filter estimator is above 5 cm. The upper limit of the matched filter estimator is dependant on other factors such as GPR probing range and the conductivity of the media. It is shown that, from Figures 5.15 and 5.25 that the thickness estimators fail abruptly. The proposed thickness estimator fails when the true thickness is greater than the thicknesses incorporated in the training data and classifier. This is the case for the experiments with real data in this chapter. Training data from layers thicker than 5 cm has not been included in these experiments as it has been shown in Sections 5.4.1 and 5.5.1 using synthetic data that the early-time feature vector is ineffective for discriminating thicker layers. The matched filter fails abruptly when the layer thickness decreases to the point where the interface reflection is masked by the clutter such as crosstalk and antenna ringdown.
Finally, averaging up to a certain SNR has the effect of minimising the variance of the estimates. This does improve the confidence interval of the proposed estimator however the reduction in the MSE is less noticeable due to the estimator bias.

For the case of a real operating or test environment (testbed or actual coal mine), the surface will ultimately be inhomogeneous and will result in variations of the dielectric properties. Under these conditions, the wave propagation velocity through the layer will vary depending on the level of variations. This will subsequently result in thickness estimator error.

Non-parallel interfaces results in varying propagation time delay from the surface to sub-surface layers. This is due to the sensor not being perpendicular to the sub-surface interface to be detected. The time delay is related to the mathematical sine of the incident angle. As the problem addressed in the thesis is defined for thin layers, the incident angle will usually be close to 90°. It is expected that non-parallel interfaces would result in insignificant change in time delay for the thin layer case and hence not affect the estimator performance significantly.

The proposed thin layer thickness estimator does not only apply to the GPR system used in this study or for estimating thin remnant coal layer thickness. The estimator still applies to other GPR systems with other centre frequencies. The operating range, however, is dependant upon the electromagnetic properties of the media being investigated and the centre frequency of the system. As the centre frequency decreases, the valid operating range increases. Further, as the dielectric constant increases, the valid operating range decreases. Additionally, there is an element of control via the choice of segmentation window length of the estimator, however there is a tradeoff between resolution and range.

Even though the proposed thickness estimator has an upper thickness limit, the method could operate in conjunction with the traditional approaches to form a complete layer thickness estimation processing scheme that covers the full range from the near-surface to the extent of the GPR range. The crossover point in Figures 5.19 and 5.19 indicate the thickness limit between thickness estimation
approaches. The near-surface interface detector proposed in Chapter 6 can govern which estimator is used in a hybrid thickness estimation scheme for a given dataset.

5.7 Summary

This chapter has presented a novel pattern recognition-based method for estimating the thickness of thin top layers using GPR. The approach was tested using both synthetic and real GPR data for two underground coal mining scenarios – coal-shale media and coal-clay media. The method was also compared with the traditional processing method for layer thickness estimation with GPR based on background subtraction and matched filter detection. The new approach is capable of estimating the thickness of thin layers in the range where the traditional approaches fail.
Chapter 6

Near-Surface Interface Detection

6.1 Introduction

Chapter 5 presented a novel method to estimate the thickness of a thin top layer. It was shown that the proposed thickness estimator is unreliable when the top layer is not thin. This chapter proposes a pattern recognition-based detector to detect sub-surface interfaces that are closer to the ground surface than other methods can operate reliably. This detector has the capability to determine if the top layer is thin. The detector has been tested using real GPR data acquired from the testbed described in Chapter 3. The antennas were surface-coupled during the acquisition of all experimental data used this chapter.

To begin the chapter, the proposed detector is presented which includes specific details of the early-time feature vector and the classifier. The experiment conducted to analyse the performance of the near-surface interface detector is also discussed. Then a discussion and the results of the experiment are presented and then a chapter summary.

6.2 Proposed Near-Surface Interface Detector

Chapter 4 showed that the proposed early-time feature vector contains valuable information for discriminating between different cases of sub-surface conditions. The objective of the near-surface interface detector is to determine if the top layer
is thin. Other information such as the bottom layer relative permittivity, conductivity or top layer thickness is not of interest. Therefore the appropriate feature vector parameters are those that provide discrimination between layer thickness less than and greater than a set threshold. The feature parameter primarily of interest for this task is the segmentation window length. A guide to selecting the window length is to inspect the images that indicate the Euclidean distances as shown in Figures 4.14 and 4.18. In the near-surface interface detection framework, these represent the distance between feature clusters for the no interface present case (bottom right corner where increasing top layer thickness converges) to other feature clusters where an interface is present. A near-surface interface is considered present if it is within a set distance threshold of the surface (ie: the top layer is thin).

Figures 4.14(b) and 4.18(b) show clear discrimination between the case of up to approximately 5 cm and Figures 4.14(c) and 4.18(c) show discrimination up to approximately 6 cm. The selected resolution is approximately 5 cm for the application of interest, therefore a window length of the order of 120 to 140 is deemed appropriate for this investigation. A window length of 128 has been selected for this task as it is in the suggested length range and is a power of two that enables the fast Fourier transform (FFT) to be used for computing the frequency domain data.

### 6.2.1 Classification

The main objective of classification in a pattern recognition system is to assign a given object to a category according to the output of the feature extractor [35].

Bayes classification is based on quantifying the tradeoffs between various decisions using probability and the costs that accompany the decisions. To use a Bayes classifier, the problem must be posed in a probabilistic manner and the probabilities must be known [35].

An optimal classifier such as the maximum likelihood classifier can easily be designed if the prior and conditional probabilities are known. In a typical situation, however, these probabilities are unknown. Methods exist to estimate the parameters of the distributions for the optimal classifier based on sampled or
training data. This is not always simple with a high dimensionality feature space. Methods such as principal component analysis (PCA) and multiple discriminant analysis (MDA) exist to decrease the dimension of feature spaces [35].

One classifier that does not require prior knowledge of parameter distributions is the artificial neural network (ANN). ANNs have become an effective classification tool in recent years [16]. Neural networks have also been used for various GPR processing problems. Some applications include detecting the presence of hyperbolic signatures in GPR images for underground pipe detection [3, 40, 98, 129] interpretation of GPR images of reinforced concrete [72] and buried landmine detection [6, 34]. In some cases, the data as either A-scans or B-scans with limited pre-processing is fed directly into the neural network classifier. The alternative is to extract features from the data which are then used as the neural network inputs. The ANN was chosen as the classifier for this task as it is fast and is capable of non-linear discrimination.

One of the most common neural network architectures is the feed-forward back-propagation having an input layer, one or more hidden layers and an output layer. The number of input units is dictated by the dimensionality of the feature vector. The number of outputs is set to the number of classes in the classification model. The exception to this rule is the two class problem which can be implemented with one unit in the output layer and target labels of -1 and 1. The number of units in the hidden layer governs the complexity of the decision boundary [35] and is generally chosen empirically. The units are connected with weights that are modified during the training process. The neural network architecture used to classify the GPR data for the near-surface interface detector is shown in Figure 6.1.

6.3 Experiment

The near-surface interface detector was tested with real GPR data acquired from the experimental testbed described in Section 3.2. Two additional detectors were implemented for comparative measures. The second method was a pattern recognition approach using features extracted from the power spectrum with a neural
network classifier. The third approach used a matched filter as a time delay estimator.

6.3.1 Processing

Pattern Recognition Detectors

The pre-processing implemented for both pattern recognition approaches were DC offset removal and sum-of-squares normalisation. Note that no background or clutter subtraction routines were utilised for these pattern recognition systems.

The parameters used for the segmentation and feature extraction stages for the early-time feature vector were a 128 sample Hamming window with a single \( a \) parameter where \( a = 1 \). Similarly, the power spectrum feature extraction stage included segmenting the pre-processed data to the first 128 samples. The power spectrum was estimated using the modified periodogram [79] utilizing the 128 point fast Fourier transform (FFT) with a Hamming window. The second and third points of the resulting modified periodogram provided sufficient discrimination between classes and hence were chosen as the feature vector for the power spectrum based pattern recognition approach.

The neural network structure used for classifying the integrated bispectrum

Figure 6.1: The neural network architecture used for classification.
and power spectrum feature vectors was the fully-connected three layer feed-forward architecture with two and four units in the input and hidden layers respectively. As the application has been cast as a two category classification problem, one unit was selected for the output layer. The activation function type was hyperbolic tangent sigmoid for the hidden layer and pure linear for the output layer.

Prior to network training, the mean and variance of the training data feature vectors were estimated. These estimates were used to standardise both the training and test data so that each feature input to the neural network was zero mean with unit variance. This ensures each feature has equal weight during the classification stage [35]. The networks were trained using the backpropagation learning algorithm with testbed data from regions 2 and 6 for the \textit{coal thickness less than 5 cm} class, and from regions 3, 4, 7, and 8 for the \textit{coal thickness greater than 5 cm} class. The training data was acquired with the antenna module positioned in the centre of each corresponding region. As the testing and training data must not be the same, the data used to test the pattern recognition systems was acquired while the antenna module was being moved over the region of interest. The category labels used during training were 1 and -1 for the \textit{coal thickness less than 5 cm} and \textit{coal thickness greater than 5 cm} classes respectively.

**Matched Filter Detector**

The matched filter cannot be implemented to detect the near-surface interface as elegantly as the pattern recognition approaches in this study. Hence, the motivation behind using the matched filter was one of time delay estimation as opposed to signal detection. The GPR signal model for layer thickness estimation in Section 5.2.1 was used to estimate the interface depth. The relative permittivity of the coal layer for the propagation velocity estimate was taken as 4.4 as given in Table 3.1. The estimate of the transmitted signal was obtained from a flat metal plate scan after subtraction of a free space calibration signal. The metal plate was 40 cm from the antenna module to ensure that the corresponding reflection was not masked by the antenna ringing.

The pre-processing stages implemented for the matched filter were background
Chapter 6. Near-Surface Interface Detection

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Coal-Clay Regions</th>
<th>Coal-Shale Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>No coal</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Coal thickness &lt; 5 cm</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Coal thickness &gt; 5 cm</td>
<td>7, 8, 9, 10</td>
<td>3, 4, 11, 12, 13, 14</td>
</tr>
</tbody>
</table>

Table 6.1: GPR testbed regions in feature clusters.

subtraction and an exponential time-varying gain function [30] to normalise for signal path losses. To allow a performance comparison between the matched filter and the pattern recognition approaches, the processed output is classified into the coal thickness less than 5 cm and coal thickness greater than 5 cm classes directly from the interface depth estimate.

6.4 Results and Discussion

The experimental results were classified into two classes according to the coal layer thickness (coal-shale and coal-clay interface depths) as less than 5 cm and greater than 5 cm with a tolerance of 0.5 cm. Figure 6.2 shows the integrated bispectral magnitude versus phase feature values for the two classes for the coal-clay and coal-shale interfaces. The data for the coal-clay interface was acquired from regions 5 to 10 while the data for the coal-shale interface was from regions 1 to 4 and 11 to 14. The regions from which the data for each cluster were acquired are shown in Table 6.1. The antenna was moved around the corresponding region during data acquisition. As a result, the feature values within the clusters are spread due to varying interface depths and the inhomogeneous layers. The confidence interval of the measured top layer thickness for each region (as shown in Table A.3 in Appendix A) indicates the layer thickness variability.

Energy reflected from an interface near the surface modifies the early-time signal of the data and hence the features in a deterministic but unknown manner. The effect on the early-time signal is more evident for the coal-clay interface than the coal-shale interface because the relative permittivity and conductivity contrast between the coal and clay is much greater than the coal and shale as seen in Figure 3.6 and Table 3.1. Hence, the clusters for the coal-clay interface
### Table 6.2: Detection and False Alarm Rates

<table>
<thead>
<tr>
<th>Interface Type</th>
<th>Feature Vector</th>
<th>Detection Rate (%)</th>
<th>False Alarm Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coal-clay</td>
<td>Integrated Bispectrum</td>
<td>94.7</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>Power Spectrum</td>
<td>80.4</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Matched Filter</td>
<td>89.3</td>
<td>0.2</td>
</tr>
<tr>
<td>Coal-shale</td>
<td>Integrated Bispectrum</td>
<td>84.1</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Power Spectrum</td>
<td>38.0</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Matched Filter</td>
<td>62.3</td>
<td>12.0</td>
</tr>
</tbody>
</table>

The data obtained for the coal thickness less than 5 cm cluster in Figure 6.2(b) was for region 2. In addition to the coal-shale interface at 25±22 mm, this region has a shale-clay interface at a total depth of 62±9 mm (from Table A.3) which is greater than 5 cm. If the reflection from the shale-clay interface is similar or greater in magnitude than the reflection from the coal-shale interface, this would affect the feature values by shifting them towards the cluster of the more dominant component.

The output of the neural network classifier is a number between -1 and 1 representing the coal thickness greater than 5 cm and coal thickness less than 5 cm classes respectively. The output of the matched filter is the interface depth estimate in centimeters. The logical choice for a simple threshold to separate the two classes is 0 for the pattern recognition approaches and 5 cm for the matched filter depth estimator. The detection and false alarm rates using these thresholds are shown in Table 6.2.

The desired performance of a detector is sometimes specified with either constant detection or false alarm rates and is usually application specific. This simplistic approach to threshold determination does not always yield optimal performance. The detector can be tweaked by changing the threshold to minimise detection errors. The detection error tradeoff (DET) curve is a useful means to compare detection schemes. The DET curves for the classifiers using the integrated bispectrum, power spectrum and matched filter for the coal-shale
interface are shown in Figure 6.3. This presents the change in detection error as the thresholds are varied.

The high false alarm rate for the matched filter at the stepped point in Figure 6.3 (miss rate \( \approx 10^2 \)) is due to the residual error from the background subtraction pre-processing stage and the inhomogeneous layers. The pattern recognition based processing does not rely on subtracting the antenna ringing or crosstalk, rather subtle changes in the signal less dominant than the ringing maintain the clusters in close proximity to the no target present case, yet provides sufficient power for detection. It can be seen that bispectral features provide better performance over the power spectrum and matched filter for detecting near-surface interfaces for this coal mining application. The early-time feature vector is thus valid for the near-surface interface detection problem. The DET curves for the coal-clay interface are not shown because the clusters are well separated as in Figure 6.2(a). The classifier has a simple task when given such well separated features as seen in the low false alarm rates in Table 6.2.
6.4. Results and Discussion

(a) Coal-clay interfaces

(b) Coal-shale and shale-clay interfaces

Figure 6.2: The bispectral magnitude and phase feature values for varying coal layer thicknesses with (a) coal-clay, (b) coal-shale and shale-clay interfaces. The clusters for no coal, coal thickness less than 5 cm and greater than 5 cm have been circled.
Figure 6.3: Detection error tradeoff (DET) curves for the bispectral features, power spectrum and the matched filter for the coal-shale interface.
The proposed near-surface interface detector can be an extremely useful pre-validation tool for the thin layer thickness estimator presented in Chapter 5. Under a combined processing strategy, the task of the near-surface interface detector is to determine if the top layer is thin, which is the case if an interface close to the surface is detected. If the top layer is thin, the thickness can subsequently be estimated using the thin layer thickness estimator. If however the top layer is not classified as being thin, the traditional approach described in Section 5.2 may be utilised to estimate the top layer thickness. The combination of the near-surface interface detector, thin layer thickness estimator and traditional thickness estimator can provide the user with the full thickness estimation range in a single GPR processing system.

6.5 Summary

This chapter has presented a pattern recognition approach for detecting the presence of interfaces that are close to the surface. The detector has been evaluated using real GPR data and compared with the classical matched filter along with the same detector using power spectral features that disregard phase information. The results of the evaluation show that the new method outperforms both the matched filter and the power spectral features for detecting near-surface interfaces.
Chapter 7

Antenna Height Estimation

7.1 Introduction

Chapter 5 presented a novel pattern recognition approach to estimate the thickness of a thin top layer using features derived from the bispectrum and a nearest neighbour classifier. This chapter adapts the methodology of the thickness estimator to an application that has practical application to the wider GPR community. That application is antenna height estimation. Additionally, the antenna height estimator can be used as a pre-validation tool for the thin layer thickness estimator to determine if the antennas are on or close to the surface.

The chapter begins with a description of the proposed method and an evaluation with both synthetic and real GPR data. The results of the experiments are then discussed and the chapter is concluded with a summary.

7.2 Problem Statement

One of the first steps of layer thickness estimation with air-coupled antennas is to estimate the antenna height above the surface. For the common application of road pavement analysis, the antenna height is constrained to within a range as specified by the equipment manufacturer. This nominated antenna height range is selected such that the first reflection is not obscured by the antenna ringing.

As described in Section 5.2.1, the antenna height for air-coupled antennas
can be estimated using equation (5.6). This is a trivial process when the ground reflection is isolated from clutter. One significant limitation, however, with having the antennas mounted high is the decreased probing range. Therefore it is useful to investigate reliable methods for estimating the antenna height when the first reflection is masked by the clutter such as antenna ringing. Figure 7.1 shows a graphical representation of the antenna height estimation task.

7.3 Proposed Antenna Height Estimator

The early-time feature vector presented in Chapter 4 has shown promise for detecting near-surface interfaces and thin layer thickness estimation with ground-coupled antennas. The principle of clutter causing problems with the early-time signal still exists when a GPR system is operated in the air-coupled configuration. Hence the thin layer thickness estimator can easily be extended to the antenna height estimation task.

7.3.1 Segmentation and Feature Extraction

The antenna height estimator consists of the same processing methods as applied to the thin layer coal thickness estimation problem. The signal is segmented using either a rectangular or Hamming window and the features are extracted.
A nearest neighbour classifier with spline for local interpolation are used for the classification phase. The training data is obtained from GPR data acquired with measured antenna heights.

### 7.3.2 Antenna Frequency Invariance

The FDTD model described in Chapter 3 was utilised to generate synthetic data for evaluating the effect of different antenna centre frequencies. The antenna centre frequency was changed so that the effect of changing antenna frequency on the feature values could be investigated. The equivalent sampling rate was also adjusted so that it is a ratio of the centre frequency. The antenna centre frequencies investigated were 500 MHz, 1 GHz and 1.5 GHz. The free-space wavelengths of these frequencies are 20 cm, 30 cm and 60 cm respectively. Figure 7.2 shows the change in feature values for a 500 MHz antenna as the antenna height varies over the range of $\lambda/20$ to $\lambda$ where $\lambda$ is the wavelength.

![Figure 7.2](image.png)

Figure 7.2: Plot showing feature values for a 500 MHz antenna as the antenna height increases up to one wavelength.

Figure 7.3 shows the feature values for 500 MHz, 1 GHz and 1.5 GHz antennas where the height is a function of wavelength and the amplitude of the raw data...
has been normalised. The figure shows that the antenna height estimator has
centre frequency invariance when the antenna height is estimated as a function of
wavelength. In a practical sense however, it is expected that the approach is not
invariant to differing antenna construction types as the antenna ringing primarily
depends on construction and design.

Figure 7.3: Plot showing feature values for a 500 MHz, 1 GHz and 1.5 GHz an-
tennas as the antenna heights increase up to one wavelength.

7.3.3 Experiment

The training data for the real GPR data experiment was acquired from an area
on the GPR testbed where the surface was coal and relatively flat. The top
layer thickness was greater than 25 cm, hence there was no near-surface interface
present. The antenna height was set and measured and data was acquired. This
process was repeated with antenna heights varying from 1 to 20 cm measured
with millimeter resolution. The height increments consisted of 1 cm from 1 to
10 cm, and 2 cm from 10 to 20 cm. Finally, the data used for the training phase
was ensemble averaged to one trace for each antenna height.
The testing dataset was acquired from various regions of the testbed at antenna heights of 2 to 20 cm with 2 cm increments. The antenna module was stationary during all data acquisition phases to ensure the antenna height could be measured reliably. The GPR antennas were mounted on a custom built timber gantry. The design and a photo of the timber gantry are shown in Figures 7.4 and 7.5 respectively. Figure 7.6 shows the GPR system in the air-coupled configuration above the testbed.

Figure 7.4: Diagram of the timber gantry constructed to mount the GPR for the air-coupled measurements.
Figure 7.5: Photo of custom built gantry and testbed for air-coupled GPR experiments.
Figure 7.6: Photo of GPR system mounted in the air-coupled configuration above the testbed.
Chapter 7. Antenna Height Estimation

Processing

The pre-processing stage of the feature extraction process included DC offset removal and a rectangular window of 128 samples was used for segmentation. The final stage before classification was feature standardisation [35] to ensure equal weighting between the integrated bispectral magnitude and phase feature vector components. Figure 7.7 shows the cubic spline, interpolated training points and true training points for the antenna height estimator.

Figure 7.7: Cubic B spline of measured and interpolated points for classifier.
7.4 Results and Discussion

Figure 7.8 shows the measured and estimated antenna heights using both the proposed method and the matched filter detector. The proposed antenna height estimator has shown an improvement over the traditional matched filter approach for estimating the antenna height in the near-field with air-coupled antennas. The near-field for these antennas is up to 6 cm from the near-field far-field threshold as mentioned in Section 3.3.1. The proposed estimator is more reliable in the near-field (first three points in Figure 7.8) than the matched filter. However the matched filter is slightly more accurate in the region from 8 cm up to 15 cm. This middle range is where the near-field effects and clutter are not significant and hence the results are as expected. In this range, the background subtraction of the free-space calibration scan is able to sufficiently remove the crosstalk and antenna ringing to reveal the ground bounce target echo.

These results are also shown in Figure 7.9, which shows the true and estimated antenna height with 99% confidence intervals. This figure clearly indicates that the bispectral feature based estimator reliably estimates the antenna height up to atleast 6 cm as opposed to the matched filter based approach which fails in this range.

The impact of averaging to reduce noise was investigated similarly as for the thin layer thickness estimator in Chapter 5. The bias, variance and MSE statistics are shown in Figures 7.10, 7.11 and 7.12 respectively. These statistics were evaluated for two antenna heights, 6 cm and 10 cm.

Figure 7.10(a) shows that the bias for the bispectral feature based height estimator is less than the matched filter estimator when the antenna height is 6 cm. The figure also shows that the bias for the matched filter at this height does not decrease as SNR increases. There appears, however, to be a relationship between bias and SNR for the bispectral feature based height estimator. As the SNR increases, the bias decreases with ensemble averaging yet increases with feature averaging. Table 4.2 indicates the same effect on bias from the two averaging approaches with synthetic data and the measured noise process.

Figure 7.10(b) indicates the converse when the antenna height is 10 cm. The
bias for the matched filter based height estimator is lower and appears to improve as SNR is increased to approximately 33 dB. However increasing the SNR beyond this point degrades the bias. The bias for the bispectral feature estimator is greater and there appears to be an insignificant relationship with SNR such that the bias changes but doesn’t converge to a significantly different value.

Figure 7.11(a) indicates that the variance for the bispectral feature based estimator is lower than the matched filter approach and there is a slight relationship with SNR. The variance improves until the SNR is approximately 27 dB.

Figure 7.11(b) indicates that the variance for the matched filter based estimator is better at the height of 10 cm and improves with ensemble averaging until the SNR is approximately 33 dB. Similar to the bias, increasing the SNR beyond 33 dB increases the variance. The figure also indicates that the variance for the bispectral feature based estimator and does not significantly improve with averaging.

The MSE statistic incorporates both bias and variance information. The trend of the MSE in Figure 7.12(a) is similar to the bias statistics in Figure 7.10(a).

Figure 7.12(b) however indicates that at an antenna height of 10 cm, the MSE
of the matched filter is similar to the variance and the MSE of the bispectral feature height estimator is similar to the bias.

Therefore the following conclusions can be made about the two antenna height estimators. Firstly, the bias, variance and MSE of each estimator is superior to the other when operating in its corresponding valid operating range. These valid operating ranges for the test GPR system and antenna are approximately 6 cm and less for the bispectral feature based estimator and greater or equal to 8 cm for the matched filter based height estimator.

Secondly, it is not indicated that the effect of averaging for these estimators improves results. Further, it is indicated that the results degrade with averaging in some cases. In particular, the results indicate that there is an optimal SNR of approximately 33 dB for the matched filter operating with an antenna height of 10 cm. Increasing the SNR beyond this optimum tends to degrade the performance of the matched filter based antenna height estimator.

Thirdly, the bispectral feature based estimator is a biased estimator as the MSE follows the trend of the bias. Also, that the matched filter based estimator is an unbiased estimator at an antenna height of 10 cm. Figure 7.9(b), however, indicates that this may not be the case for antenna heights greater than 10 cm as the height appears to be underestimated in this range. Therefore, it is concluded that both estimators on average are biased estimators. It is likely that this bias is due systematic causes such as non-linearities in the GPR sampling system. Another likely cause is height measurement variability during the experiment.

The upper limit for the bispectral feature based antenna height estimator is shown to be approximately 7 cm for the GPR system used in this study. The wavelength for the antenna (centre frequency of 1.4 GHz) is 21.4 cm. As a function of wavelength, this valid range is approximately up to $\lambda/3$. It has also been shown that the proposed antenna height estimator is relatively invariant to centre frequency when considered as a function of wavelength. Therefore it is concluded that the novel antenna height estimator proposed in this chapter can be applied to many different GPR antennas for estimating antenna height up to $\lambda/3$ for the given antenna.
Figure 7.9: True and estimated antenna heights using (a) proposed and (b) matched filter based antenna height estimators. The error bars represent the 99% confidence interval.
Figure 7.10: Plots show the bias for the bispectral feature and matched filter based antenna height estimators as the SNR is increased through averaging. The bias is shown for antenna heights of (a) 6 cm and (b) 10 cm.
Figure 7.11: Plots show the variance for the bispectral feature and matched filter based antenna height estimators as the SNR is increased through averaging. The variance is shown for antenna heights of (a) 6 cm and (b) 10 cm.
7.4. Results and Discussion

Figure 7.12: Plots show the MSE for the bispectral feature and matched filter based antenna height estimators as the SNR is increased through averaging. The MSE is shown for antenna heights of (a) 6 cm and (b) 10 cm.
7.5 Summary

In this chapter, a novel pattern recognition-based method for estimating the height of air-coupled GPR antennas when the antennas are close to the surface has been presented. It has been shown that the proposed approach is reliable for antenna heights up to one third of the wavelength of the given antenna. Traditional height estimators such as one based on pre-processing and the matched filter are unreliable when the height is less than one third of the wavelength, but operate reasonably successfully when the height is greater than this threshold.

The significance of this outcome is that air-coupled antennas can be operated closer to the ground surface to increase probing range yet still obtain quantitative information about the height of the antenna (which is necessary for target depth estimation processing routines) rather than keep the antenna further away as required for traditional antenna height estimation approaches.

The proposed antenna height estimator has been tested with synthetic data and validated with real data acquired from the GPR system described in Chapter 3.
Chapter 8

Summary and Future Research

The need for layer thickness estimation in underground coal mining was presented and a review of coal-rock interface sensing methods given. A review of GPR was also presented which discussed the common problems associated with thickness estimation of thin layers.

The primary source of data used in the experiments was described. This included the design and construction of a testbed and measurement of dielectric parameters for validating the proposed processing strategies, the details of the low power GPR system used for real data acquisition and the FDTD method for synthetic data generation.

The early-time feature vector that is able to characterise the early-time signal of GPR data was proposed. This feature vector is central to the thin layer thickness estimator, near-surface interface detector and antenna height estimator. A review of higher order spectra with special attention given to the bispectrum was also presented including a review of pattern recognition work applied to GPR.

A novel thin layer thickness estimation processing method was proposed. The technique is a pattern recognition approach utilising the early-time feature vector and a nearest neighbour classifier. This method was compared with the traditional GPR layer thickness estimation strategy that is based on the matched filter. It was shown that the proposed thin layer thickness estimator is superior to the traditional layer thickness approach when the top layer is thin. Conversely however, the traditional approach is more effective when the top layer is not thin. There is an upper top layer thickness limit where the proposed thin layer
thickness estimator fails abruptly. This limit is the boundary where the proposed approach is unsuccessful and the traditional approach is recommended.

The proposed thin layer thickness estimator does not only apply to the GPR system used in this study or for estimating remnant coal layer thickness. The proposed methods can be applied to a wide variety of commercial impulse GPR systems because the problem of clutter such as antenna ringing and crosstalk are common to all GPR systems.

The processing method to detect near-surface interfaces was introduced and evaluated. The proposed approach as well as a matched filter detector were tested and quantitatively analysed using real GPR data. The results show that the proposed near-surface interface detector based on the early-time feature vector is superior to the matched filter detector and features based on the power spectrum. This is evident from the DET curve in Figure 6.3.

The coal in the testbed is unconsolidated but was compacted very well when the testbed layers were formed. It is expected that the performance of the thin layer thickness estimator and near-surface interface detector will improve when tested with data acquired from coal in-situ such as an underground coal mine. This is because the data from the testbed may have additional clutter which resulted from the electromagnetic energy reflecting from small airgaps that may be present beneath the testbed surface.

The thin layer thickness estimator proposed in Chapter 5 has an upper limit to which the thickness can be estimated. Beyond this thickness, the proposed estimator fails and traditional layer thickness estimation approaches are more successful. Therefore a system that utilises the proposed thickness estimator must have information as to whether the layer is thin. This information can be provided by the near-surface interface detector. A hybrid system incorporating these two tools can operate to estimate the thickness of a thin top layer in a robust fashion using GPR.

A novel method for estimating the height of air-coupled antennas above the ground surface was described. Results show that the proposed method is more robust than traditional matched filter based approaches when the antenna height is less than one third of the electromagnetic wavelength of the antenna. The
antenna height estimator can be used to either determine if the antenna height is close to the surface such that a classifier trained for the thin layer thickness estimation task is valid for a given dataset. Alternatively, the antenna height estimator could be used in the road pavement evaluation application with the air-coupled antennas close to the surface which will increase the probing range of these systems.

8.1 Directions for Future Research

Some ideas for future research include the following:

1. Trial the proposed signal processing methods in an underground mining environment. One key challenge with this is the need to collect real data from an underground coal mine under known conditions to train the classifiers.

2. Development of a full 3D FDTD model of GPR antenna module and associated equipment that can be used to train the classifiers. This will eliminate the need to physically excavate the surface to measure the top layer thickness.

3. Evaluate the early-time feature vector and novel approaches proposed in this dissertation with other non-coal surface materials.

4. Develop enhancements to the antenna height estimator that involves other top layer media. Specifically, this would be to investigate an approach that is invariant to events such as other top layer surface types, near-surface interface reflections and surface roughness. Additionally, methods to estimate both the antenna height and the thickness of a thin top layer simultaneously with air-coupled antennas can be investigated.

5. Investigate other important signal processing aspects of the estimators such as error models, Cramer-Rao lower bounds, sensitivity analysis and mismatch between training and testing data.
Appendix A

Testbed Measurements

This appendix presents the measured layer thickness of the testbed described in Chapter 3. A method to estimate thin layer thickness was proposed and evaluated in Chapter 5. Table A.1 presents the measured layer thickness used to evaluate the estimator with the coal-clay interface. Table A.2 presents the measured layer thickness used to evaluate the estimator with the coal-shale interface.

Chapter 6 proposed a method for detecting the presence of interfaces close to the surface. The data for the evaluation of this method was acquired from the 14 testbed regions. The measured coal and shale layer thickness and tolerances, and clay interface depth for each region are shown in Table A.3. The tolerance measurements were calculated from the measured thicknesses/depths for each region with 95% confidence covering 90% of each region surface [45].
### Table A.1: Measured coal layer thickness for the coal-clay interface evaluation of the thin layer thickness estimator.

<table>
<thead>
<tr>
<th>Position (cm)</th>
<th>Coal Thickness (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>51</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
</tr>
<tr>
<td>10</td>
<td>51</td>
</tr>
<tr>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>20</td>
<td>34</td>
</tr>
<tr>
<td>25</td>
<td>27</td>
</tr>
<tr>
<td>30</td>
<td>21</td>
</tr>
<tr>
<td>35</td>
<td>14</td>
</tr>
<tr>
<td>40</td>
<td>11</td>
</tr>
<tr>
<td>45</td>
<td>6</td>
</tr>
<tr>
<td>50</td>
<td>32</td>
</tr>
<tr>
<td>55</td>
<td>26</td>
</tr>
<tr>
<td>60</td>
<td>16</td>
</tr>
<tr>
<td>65</td>
<td>18</td>
</tr>
<tr>
<td>70</td>
<td>18</td>
</tr>
<tr>
<td>75</td>
<td>20</td>
</tr>
<tr>
<td>80</td>
<td>19</td>
</tr>
<tr>
<td>85</td>
<td>25</td>
</tr>
<tr>
<td>90</td>
<td>19</td>
</tr>
<tr>
<td>95</td>
<td>19</td>
</tr>
<tr>
<td>100</td>
<td>22</td>
</tr>
<tr>
<td>$x$ Position (cm)</td>
<td>Coal Thickness (mm)</td>
</tr>
<tr>
<td>------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>105</td>
<td>13</td>
</tr>
<tr>
<td>110</td>
<td>29</td>
</tr>
<tr>
<td>115</td>
<td>27</td>
</tr>
<tr>
<td>120</td>
<td>23</td>
</tr>
<tr>
<td>125</td>
<td>21</td>
</tr>
<tr>
<td>130</td>
<td>21</td>
</tr>
<tr>
<td>135</td>
<td>26</td>
</tr>
<tr>
<td>140</td>
<td>36</td>
</tr>
<tr>
<td>145</td>
<td>30</td>
</tr>
<tr>
<td>150</td>
<td>28</td>
</tr>
<tr>
<td>155</td>
<td>33</td>
</tr>
<tr>
<td>160</td>
<td>33</td>
</tr>
<tr>
<td>165</td>
<td>31</td>
</tr>
<tr>
<td>170</td>
<td>28</td>
</tr>
<tr>
<td>175</td>
<td>31</td>
</tr>
<tr>
<td>180</td>
<td>35</td>
</tr>
<tr>
<td>185</td>
<td>27</td>
</tr>
<tr>
<td>190</td>
<td>22</td>
</tr>
<tr>
<td>195</td>
<td>23</td>
</tr>
<tr>
<td>200</td>
<td>17</td>
</tr>
<tr>
<td>205</td>
<td>12</td>
</tr>
<tr>
<td>210</td>
<td>6</td>
</tr>
<tr>
<td>215</td>
<td>13</td>
</tr>
<tr>
<td>220</td>
<td>15</td>
</tr>
<tr>
<td>225</td>
<td>22</td>
</tr>
<tr>
<td>230</td>
<td>22</td>
</tr>
<tr>
<td>235</td>
<td>20</td>
</tr>
<tr>
<td>240</td>
<td>20</td>
</tr>
</tbody>
</table>

Table A.2: Measured coal and shale layer thickness for the coal-shale interface evaluation of the thin layer thickness estimator.
<table>
<thead>
<tr>
<th>Region</th>
<th>Coal</th>
<th>Shale</th>
<th>Clay (depth)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>42 ± 21 mm</td>
<td>42 ± 21 mm</td>
</tr>
<tr>
<td>2</td>
<td>26 ± 22 mm</td>
<td>37 ± 22 mm</td>
<td>62 ± 9 mm</td>
</tr>
<tr>
<td>3</td>
<td>63 ± 23 mm</td>
<td>37 ± 21 mm</td>
<td>101 ± 26 mm</td>
</tr>
<tr>
<td>4</td>
<td>112 ± 31 mm</td>
<td>54 ± 56 mm</td>
<td>166 ± 34 mm</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>0 mm</td>
</tr>
<tr>
<td>6</td>
<td>43 ± 13 mm</td>
<td>-</td>
<td>43 ± 13 mm</td>
</tr>
<tr>
<td>7</td>
<td>69 ± 21 mm</td>
<td>-</td>
<td>69 ± 21 mm</td>
</tr>
<tr>
<td>8</td>
<td>119 ± 18 mm</td>
<td>-</td>
<td>119 ± 18 mm</td>
</tr>
<tr>
<td>9</td>
<td>220 ± 26 mm</td>
<td>-</td>
<td>220 ± 26 mm</td>
</tr>
<tr>
<td>10</td>
<td>309 ± 21 mm</td>
<td>-</td>
<td>309 ± 21 mm</td>
</tr>
<tr>
<td>11</td>
<td>117 ± 14 mm</td>
<td>42 ± 36 mm</td>
<td>159 ± 31 mm</td>
</tr>
<tr>
<td>12</td>
<td>211 ± 36 mm</td>
<td>47 ± 14 mm</td>
<td>258 ± 29 mm</td>
</tr>
<tr>
<td>13</td>
<td>293 ± 65 mm</td>
<td>40 ± 62 mm</td>
<td>334 ± 44 mm</td>
</tr>
<tr>
<td>14</td>
<td>338 ± 45 mm</td>
<td>158 ± 54 mm</td>
<td>496 ± 27 mm</td>
</tr>
<tr>
<td>15</td>
<td>Varying</td>
<td>51 ± 61 mm</td>
<td>Varying</td>
</tr>
</tbody>
</table>

Table A.3: Measured thickness for coal and shale layers, and clay interface depth.
Appendix B

Noise Analysis of GPR System

The radar signal model presented in Section 5.2.1 treated the received signal as the superposition of target echoes plus noise. There are various noise sources in impulse GPR systems that corrupt the raw data. These sources are traditionally categorised into two types, namely clutter and additive noise. Daniels [30] defined clutter as the “signals that are unrelated to the target scattering characteristics but occur in the same sample time window and have similar spectral characteristics to the target wavelet”. At a given spatial position with stationary surroundings, the clutter components are repeatable. Typical sources of clutter in GPR are the crosstalk, antenna ringdown, target echo multiples and echoes from non-target reflectors.

Unlike the clutter components, additive noise is random and hence not repeatable. Sources of the additive noise include quantisation noise, timing/amplitude jitter and thermal noise. Quantisation noise results from the sampling process where the data is converted from a continuous analog signal to discrete quantised levels [99]. Timing and amplitude jitter is the consequence of instability in the sequential sampling system [30]. Thermal noise is generated by the thermal agitation of electrons in any electrical conductor [80].

For the purpose of evaluating if the common assumption that the additive noise is normally distributed as assumed in the radar signal model of Section 5.2.1, it is useful to investigate the statistical properties of the additive noise component in the GPR data. Data for this investigation was acquired with the GPR antennas in both the free-space configuration and surface coupled on a coal
layer. The additive noise process was estimated by subtracting the ensemble average of the raw test data and hence does not include clutter components.

Figure B.1: Figure shows the residual noise after mean has been subtracted from a free-space scan.

Figure B.1 shows the residual noise after the ensemble average has been subtracted from the data. The noise at the first time segment of the process (first 100 samples) is significantly higher compared with the remaining samples. This is the timeframe of the early-time signal. This region of higher noise power is also clearly visible from the variance as shown in Figure B.2.

The Bera and Jarque normality test [28, 47] takes advantage of the fact that the odd moments above two are zero for the normal distribution. The third and fourth central moments, also known as skew and kurtosis respectively, are 0 and close to 3 for a normally distributed random variable with a sample variance of 1 [28]. The skewness of a distribution is a measure of the lack of symmetry around the mean and kurtosis is influenced by the peakness of the distribution and thickness in the tails [28, 47]. The skewness and kurtosis for unit sample variance are described mathematically below.

For the sampled radar signal \( r_m(t) \) with \( M \) realisations, the \( n^{th} \) central moment
Figure B.2: Figure shows the variance of the additive noise in the raw GPR data with the antennas in (a) free-space, and (b) on the coal surface.

at time $t$ is

$$\mu_n(t) = \frac{1}{M} \sum_{m=1}^{M} (r_m(t) - \bar{r}(t))^n$$  \hfill (B.1)

where $m = 1, \ldots, M$ is the realisation index and $\bar{r}(t)$ is the mean of $r_m(t)$ over all realisations. The measure of skewness is given by

$$\sqrt{b_1(t)} = \frac{\mu_3(t)}{\sigma^3(t)}$$  \hfill (B.2)

and the measure of kurtosis is given by

$$b_2(t) = \frac{\mu_4(t)}{\sigma^4(t)}$$  \hfill (B.3)

where $\mu_3(t)$ and $\mu_4(t)$ are the third and fourth central moments respectively and $\sigma^2(t)$ is the variance. The Bera-Jarque test is a joint test of whether or not the estimates of the skewness and kurtosis differ significantly to those values for the normal distribution. The test statistic for the Bera-Jarque test is an asymmetric chi-squared distribution with two degrees of freedom ($\chi^2_2$) [47]

$$\lambda(t) = M \left( \frac{b_1(t)}{6} + \frac{[b_2(t) - 3]^2}{24} \right)$$

$$= M \left( \frac{\mu_3^2(t)}{6\sigma^6(t)} + \frac{[\mu_4(t) - 3\sigma^4(t)]^2}{24\sigma^8(t)} \right)$$  \hfill (B.4)

The statistical properties of the additive noise process are shown in Figure B.3. The skewness is shown in (a,b), the kurtosis is shown in (c,d) and the Bera-Jarque test statistic and critical value are shown in (e,f). The null hypothesis
of the distribution is normal is accepted when the test statistic is less than the critical value. The critical value for a chi-squared distribution with two degrees of freedom $\chi^2_2$ and $\alpha=0.05$ is 5.991 [47].

The results of the hypothesis test indicate the presence of two types of noise in the data. Firstly, it is likely that the region of the first 200 samples of the data does not have a normal distribution. This is where the test statistic (refer to Figure B.3 (e,f)) is greater than the critical value. This non-Gaussian noise is due to timing and amplitude jitter, and exists in the early-time segment of the data. Timing and amplitude jitter is more noticeable in the region where the signal is changing rapidly [30].

The timing and amplitude jitter is the cause of both the primary and secondary peaks in Figure B.2. The primary peak at approximately sample 30 is the maximum point in the early-time signal and is at approximately 0.25 ns in Figure 3.8. Similarly, the secondary peak at sample 70 is the minimum point at approximately 0.7 ns in Figure 3.8.

Secondly, it is likely that the region after the early-time signal is normally distributed as the test statistic is close to the critical value.

The level of quantisation noise was investigated. The sampling parameters presented in section 3.3.1 were used to calculate the quantisation noise as zero mean with a variance of $1.24 \times 10^{-7}$ [80, 99]. As the quantisation noise variance stated here is several orders of magnitude smaller than the other noise variance shown in Figure B.2, quantisation noise effects are considered insignificant.

It is concluded that the noise in the first part of the GPR trace does not possess a Gaussian distribution whereas it is likely that the remainder of the trace does. This has implications in terms of motivating the use of alternative methods for signal detection and estimation of the early-time signal as presented in Chapters 4, 5, 6 and 7.
Figure B.3: Skewness, kurtosis and test statistic of the additive noise in the raw GPR data with the antennas in (a) free-space, and (b) on the coal surface.
Appendix C

Finite-Difference Time-Domain Update Equations

C.1 Introduction

This appendix presents the 2D FDTD update equations implemented to generate the synthetic data used in this dissertation. This is not intended as a tutorial on the FDTD method as there are many references in the literature that cover the FDTD method thoroughly such as the book by Taflove [111] and other works such as [11, 107].

C.2 2D FDTD Update Equations

The transverse magnetic (TM) FDTD update equations in 2D are [107]

\[
H_x|^{n+1/2}_{i,j} = \left(1 - \frac{\rho_{i,j}^l \Delta t}{2 \mu_{i,j}}\right) H_x|^{n-1/2}_{i,j} + \left(\frac{\Delta t}{\mu_{i,j}}\right) \cdot \left(\frac{1}{1 + \frac{\rho_{i,j}^l \Delta t}{2 \mu_{i,j}}} \right) \cdot \left(-\frac{E_z|^{n}_{i+1/2,j} - E_z|^{n}_{i-1/2,j}}{\Delta y}\right)
\]

\[
H_y|^{n+1/2}_{i,j} = \left(1 - \frac{\rho_{i,j}^l \Delta t}{2 \mu_{i,j}}\right) H_y|^{n-1/2}_{i,j} + \left(\frac{\Delta t}{\mu_{i,j}}\right) \cdot \left(\frac{1}{1 + \frac{\rho_{i,j}^l \Delta t}{2 \mu_{i,j}}} \right) \cdot \left(-\frac{E_z|^{n}_{i+1/2,j} - E_z|^{n}_{i-1/2,j}}{\Delta x}\right)
\]
\[ E_{z}^{n+1}_{i,j} = \left( 1 - \frac{\sigma_{ij} \Delta t}{2 \varepsilon_{i,j}} \right) E_{z}^{n}_{i,j} + \left( \frac{\Delta t}{\varepsilon_{i,j} \Delta x} \right) \left( \frac{H_{y}^{n+1/2}_{i+1/2,j} - H_{y}^{n+1/2}_{i-1/2,j}}{\Delta x} - \frac{H_{x}^{n+1/2}_{i,j+1/2} - H_{x}^{n+1/2}_{i,j-1/2}}{\Delta y} \right) \]

where \( H_{x} \) and \( H_{y} \) are the magnetic field components in the \( x \) and \( y \) directions respectively, \( E_{z} \) is the electric field component in the \( z \) direction, \( \rho' \) is an equivalent magnetic resistivity in Ohms per metre, \( \sigma \) is the electric conductivity in Siemens per metre, \( \mu \) is the magnetic permeability in Henrys per metre, \( \varepsilon \) is the electric permittivity in Farads per metre, \( i \) and \( j \) are position indices in the FDTD grid, \( \Delta x \) and \( \Delta y \) are the grid spacing in the \( x \) and \( y \) directions respectively, \( \Delta t \) is the time step and \( n \) is the time step index.

### C.3 Dispersion Model Update Equations

To incorporate dispersion in the FDTD implementation using the Debye model, the update equation of the electric field \( E_{z} \) is modified as follows [11]

\[ E_{z}^{n+1}_{i,j} = \left( 2 \varepsilon_{0} \varepsilon_{r_{\infty}} - \sigma_{0} \Delta t \right) \left( 2 \varepsilon_{0} \varepsilon_{r_{\infty}} + 2 \varepsilon_{0} \chi_{0} + \sigma_{0} \Delta t \right) E_{z}^{n}_{i,j} + \left( 2 \Delta t \right) \left( \frac{2 \varepsilon_{0} \varepsilon_{r_{\infty}} + 2 \varepsilon_{0} \chi_{0} + \sigma_{0} \Delta t}{2 \varepsilon_{0} \varepsilon_{r_{\infty}} + 2 \varepsilon_{0} \chi_{0} + \sigma_{0} \Delta t} \right) \Psi_{z}^{n}_{i,j} + \left( \frac{H_{y}^{n+1/2}_{i+1/2,j} - H_{y}^{n+1/2}_{i-1/2,j}}{\Delta x} - \frac{H_{x}^{n+1/2}_{i,j+1/2} - H_{x}^{n+1/2}_{i,j-1/2}}{\Delta y} \right) \Psi_{z}^{n}_{i,j} \]

where \( \varepsilon_{0} \) is the permittivity of free space \( (8.854 \times 10^{-12} \text{ F/m}) \) [92], \( \varepsilon_{r_{\infty}} \) is the static frequency relative permittivity, \( \varepsilon_{r_{\infty}} \) is the infinite frequency relative permittivity, \( \tau \) is the relaxation time and \( \sigma_{0} \) is the low frequency conductivity as described in Section 3.4.1, and

\[ \Psi_{z}^{n+1}_{i,j} = \Delta \chi_{0} E_{z}^{n+1}_{i,j} + (e^{-\Delta t/\tau}) \Psi_{z}^{n}_{i,j} \]

\[ \chi_{0} = (\varepsilon_{r_{\infty}} - \varepsilon_{r_{\infty}}) \left( 1 - e^{-\Delta t/\tau} \right) \]

\[ \Delta \chi_{0} = (\varepsilon_{r_{\infty}} - \varepsilon_{r_{\infty}}) \left( 1 - e^{-\Delta t/\tau} \right)^{2} \]
C.4 Courant Stability Criterion

The Courant stability criterion imposes an upper limit of the time step $\Delta t$ to ensure numerical stability. The time step limit is dependant upon the cell spacing and propagation velocity in the medium. The time step limit for a 2D FDTD grid with $\Delta x \neq \Delta y$ is [11]

$$\Delta t \leq \frac{1}{v \sqrt{\frac{1}{(\Delta x)^2} + \frac{1}{(\Delta y)^2}}}$$

where $v$ is the maximum electromagnetic wave propagation velocity in the grid.

For a 2D FDTD grid with $\Delta x = \Delta y = \Delta$, the time step limit simplifies to

$$\Delta t \leq \frac{\Delta}{v\sqrt{2}}$$

C.5 Absorbing Boundary Conditions

The second order absorbing boundary conditions proposed by Mur [69] were implemented to minimise artificial reflections from the grid boundaries. The grid was implemented with the electric field component ($E_z$) on the outer edges of the grid. The update equation for the $x = 0$ boundary is

$$E_z|^{n+1}_{0,j} = -E_z|^{n-1}_{1,j} + \frac{c\Delta t - \Delta x}{c\Delta t + \Delta x} \left( E_z|^{n+1}_{1,j} + E_z|^{n-1}_{0,j} \right) + \frac{2\Delta x}{c\Delta t + \Delta x} \left( E_z|^{n}_{0,j} + E_z|^{n}_{1,j} \right) + \frac{(c\Delta t)^2 \Delta x}{2\Delta y^2 (c\Delta t + \Delta x)} \cdot \left( E_z|^{n}_{0,j+1} - 2E_z|^{n}_{0,j} + E_z|^{n}_{0,j-1} + E_z|^{n+1}_{1,j+1} - 2E_z|^{n}_{1,j} + E_z|^{n}_{1,j-1} \right)$$

The update equation for the $y = 0$ boundary is

$$E_z|^{n+1}_{i,0} = -E_z|^{n-1}_{i,1} + \frac{c\Delta t - \Delta y}{c\Delta t + \Delta y} \left( E_z|^{n+1}_{i,1} + E_z|^{n-1}_{i,0} \right) + \frac{2\Delta y}{c\Delta t + \Delta y} \left( E_z|^{n}_{i,0} + E_z|^{n}_{i,1} \right) + \frac{(c\Delta t)^2 \Delta y}{2\Delta x^2 (c\Delta t + \Delta y)} \cdot \left( E_z|^{n}_{i+1,0} - 2E_z|^{n}_{i,0} + E_z|^{n}_{i-1,0} + E_z|^{n+1}_{i+1,1} - 2E_z|^{n}_{i,1} + E_z|^{n}_{i-1,1} \right)$$
The update equation for the $x = h_x$ boundary is

\[
E_{z|hx,j}^{n+1} = - E_{z|hx-1,j}^{n-1} + \frac{c \Delta t - \Delta x}{c \Delta t + \Delta x} \left( E_{z|hx-1,j}^{n+1} + E_{z|hx-1,j}^{n-1} \right) + \frac{2 \Delta x}{c \Delta t + \Delta x} \left( E_{z|hx,j}^{n} + E_{z|hx-1,j}^{n} \right) + \frac{(c \Delta t)^2 \Delta x}{2 \Delta y^2 (c \Delta t + \Delta x)} \cdot \left\{ \left[ E_{z|hx,j+1}^{n} + E_{z|hx,j-1}^{n} + E_{z|hx-1,j+1}^{n} + E_{z|hx-1,j-1}^{n} \right] - 2 \left[ E_{z|hx,j}^{n} + E_{z|hx-1,j}^{n} \right] \right\}
\]

where $h_x$ is the grid dimension along the $x$ co-ordinate. The update equation for the $y = h_y$ boundary is

\[
E_{z|i,hy}^{n+1} = - E_{z|i,hy-1}^{n-1} + \frac{c \Delta t - \Delta y}{c \Delta t + \Delta y} \left( E_{z|i,hy-1}^{n+1} + E_{z|i,hy-1}^{n-1} \right) + \frac{2 \Delta y}{c \Delta t + \Delta y} \left( E_{z|i,hy}^{n} + E_{z|i,hy-1}^{n} \right) + \frac{(c \Delta t)^2 \Delta y}{2 \Delta x^2 (c \Delta t + \Delta y)} \cdot \left\{ \left[ E_{z|i+1,hy}^{n} + E_{z|i-1,hy}^{n} + E_{z|i+1,hy-1}^{n} + E_{z|i-1,hy-1}^{n} \right] - 2 \left[ E_{z|i,hy}^{n} + E_{z|i,hy-1}^{n} \right] \right\}
\]

where $h_y$ is the grid dimension along the $y$ co-ordinate.
Bibliography


