Multiresolution Analysis of Ultrasound Images of the Prostate

Fangwei Zhao
B.E., M.Sc.

This thesis is presented for the degree of Doctor of Philosophy of
The University of Western Australia

Centre for Intelligent Information Processing Systems
School of Electrical, Electronic and Computer Engineering
The University of Western Australia
Crawley, WA 6009
Australia

November 2003
"Oh, yeah? Well, my dad's prostate is more enlarged than your dad's!"

Abstract

Transrectal ultrasound (TRUS) has become the urologist’s primary tool for diagnosing and staging prostate cancer due to its real-time and non-invasive nature, low cost, and minimal discomfort. However, the interpretation of a prostate ultrasound image depends critically on the experience and expertise of a urologist and is still difficult and subjective. To overcome the subjective interpretation and facilitate objective diagnosis, computer aided analysis of ultrasound images of the prostate would be very helpful. Computer aided analysis of images may improve diagnostic accuracy by providing a more reproducible interpretation of the images.

This thesis is an attempt to address several key elements of computer aided analysis of ultrasound images of the prostate. Specifically, it addresses the following tasks:

1. modelling B-mode ultrasound image formation and statistical properties;
2. reducing ultrasound speckle; and
3. extracting prostate contour.

Speckle refers to the granular appearance that compromises the image quality and resolution in optics, synthetic aperture radar (SAR), and ultrasound. Due to the existence of speckle the appearance of a B-mode ultrasound image does not necessarily relate to the internal structure of the object being scanned. A computer simulation of B-mode ultrasound imaging is presented, which not only provides an insight into the nature of speckle, but also a viable test-bed for any ultrasound speckle reduction methods. Motivated by analysis of the statistical properties of the simulated images, the generalised Fisher-Tippett distribution is empirically proposed to analyse statistical properties of ultrasound images of the prostate.

A speckle reduction scheme is then presented, which is based on Mallat and Zhong’s dyadic wavelet transform (MZDWT) and modelling statistical properties of the wavelet coefficients and exploiting their inter-scale correlation. Specifically, the squared modulus of the component wavelet coefficients are modelled as a two-state Gamma mixture. Inter-scale correlation is exploited by taking the harmonic mean of the posterior probability functions, which are derived from the Gamma mixture. This noise reduction scheme is applied to both simulated and real ultrasound images, and its performance is quite satisfactory in that the important features of the original noise corrupted image are preserved
while most of the speckle noise is removed successfully. It is also evaluated both qualitatively and quantitatively by comparing it with median, Wiener, and Lee filters, and the results revealed that it surpasses all these filters.

A novel contour extraction scheme (CES), which fuses MZDWT and snakes, is proposed on the basis of multiresolution analysis (MRA). Extraction of the prostate contour is placed in a multi-scale framework provided by MZDWT. Specifically, the external potential functions of the snake are designated as the modulus of the wavelet coefficients at different scales, and thus are “switchable”. Such a multi-scale snake, which deforms and migrates from coarse to fine scales, eventually extracts the contour of the prostate. The CES is evaluated quantitatively using both distance-based and area-based criteria. It surpasses the manual outline with respect to the distance-based metrics and is comparable with the manual outline in terms of the area-based criterion. This indicates that the CES is reliable and could be used as a highly accurate alternative reference for the tedious, subjective and hardly reproducible manual outlining process.
Contents

Abstract i

Acknowledgements xv

1 Introduction 1

1.1 Computerised Analysis of Medical Images 2

1.2 Structure of Thesis 4

1.3 Outline of Thesis 5

1.3.1 Ultrasound Image Modelling 6

1.3.2 Ultrasound Speckle Reduction 6

1.3.3 Prostate Contour Extraction 7

2 Ultrasound of the Prostate: The Background 9

2.1 The Prostate 9

2.1.1 Function of the Prostate 9

2.1.2 Prostatic Disease 10

2.1.3 Diagnosis of Prostate Cancer 13

2.1.4 Treatment of Prostate Cancer 16

2.1.5 Summary 17

2.2 B-mode Ultrasound Imaging 17

2.2.1 Pre-processing 18

2.2.2 Noise Suppression by Thresholding 19

2.2.3 Time Gain Compensation (TGC) 19

2.2.4 Demodulation (Rectification and Envelope Detection) 20

2.2.5 Digital Scan Conversion 20

2.2.6 Post-processing 20

2.2.7 Summary 20

2.3 The Proscan Plus Imaging System 21

2.4 Appearance of Ultrasound Images of the Prostate 22

2.5 Sample Ultrasound Images of the Prostate 25
3 Speckle Distribution Models: An Overview

3.1 Overview of Speckle Distribution Models

3.1.1 The Rayleigh Distribution

3.1.2 The Rician Distribution

3.1.3 The K Distribution

3.1.4 The Fisher-Tippett Distribution

3.1.5 The Nakagami Distribution

3.1.6 The Weibull Distribution

3.1.7 The Generalised Gamma Distribution

3.1.8 The Generalised Fisher-Tippett Distribution

3.2 Parameter Estimation

3.2.1 Maximum Likelihood Estimation (MLE)

3.2.2 Nonlinear Curve Fitting

3.2.3 The Method of Moments and Initial Estimation of Parameters

3.3 Goodness-of-Fit Test

3.3.1 The Chi-Squared Goodness-of-Fit Test

3.3.2 The Kolmogorov-Smirnov One-Sample Test

3.4 Summary

4 Computer Simulation of B-mode Ultrasound Imaging

4.1 Hypothetical B-mode Ultrasonic Imaging System

4.1.1 Imaging System Modelling

4.1.2 Modelling the Tissue

4.1.3 Envelope Detection

4.1.4 Logarithmic Compression

4.1.5 Scatterer Density

4.2 Simulated Images and Statistical Analysis

4.2.1 Homogeneous Phantoms

4.2.2 Homogeneous Phantoms with Effect of Scatterer Density $d_s$

4.2.3 Homogeneous Phantoms with Effect of Logarithmic Compression

4.2.4 Phantoms with Coherent Components

4.3 Statistical Analysis of Real Ultrasound Images

4.4 Conclusion

5 Speckle Noise Reduction: A Survey

5.1 System Level Speckle Noise Reduction

5.1.1 Compound Scanning

5.1.2 Seggie's Phased-Based Deterministic Approach

5.1.3 Hokland's Harmonic Imaging

5.1.4 Comments on System Level Speckle Reduction

5.2 Image Level Speckle Noise Reduction
CONTENTS

5.2.1 Classical Adaptive Filtering Methods..................................................... 93
5.2.2 Wavelet Methods.................................................................................. 99
5.2.3 Speckle Reduction by Geometry Driven Diffusion................................. 105
5.2.4 Other Methods..................................................................................... 107
5.3 Discussion................................................................................................. 108

6 Speckle Reduction Using the Dyadic Wavelet Transform .......................... 111
6.1 Introduction............................................................................................... 111
6.2 Wavelets and the Dyadic Wavelet Transform......................................... 112
  6.2.1 Wavelets and the Continuous Wavelet Transform............................ 112
  6.2.2 Multiresolution Analysis...................................................................... 115
  6.2.3 Biorthogonal Wavelets......................................................................... 123
  6.2.4 Mallat and Zhong’s Dyadic Wavelet Transform.............................. 125
6.3 Noise Reduction Scheme in the Wavelet Transform Domain.................... 134
  6.3.1 Modelling Statistical Properties of the Wavelet Coefficients........... 135
  6.3.2 Wavelet Coefficients Shrinkage Using Inter-scale Correlation.......... 142
  6.3.3 Summary of the Noise Reduction Scheme........................................ 146
6.4 Results and Validation............................................................................. 146
  6.4.1 Qualitative Evaluation....................................................................... 148
  6.4.2 Quantitative Validation...................................................................... 150
  6.4.3 Application to Real Ultrasound Images........................................... 151
6.5 Conclusion................................................................................................. 152

7 Contour Detection Using the Laplacian of Gaussian Operator.................... 157
7.1 The Laplacian of Gaussian (LoG) Operator............................................ 157
7.2 Contour Detection Based on the LoG Operator....................................... 160
7.3 Results and Discussion............................................................................ 161
7.4 Conclusion................................................................................................. 163

8 Delineation of Prostate Contours: A Critical Review............................... 165
8.1 Introduction............................................................................................... 165
8.2 Pixel or Texture Based Strategies........................................................... 171
  8.2.1 Prater’s Feed-Forward Neural Network........................................... 171
  8.2.2 Richard’s Texture Based Segmentation.......................................... 171
  8.2.3 Comments on Pixel or Texture Based Segmentation...................... 171
8.3 Edge Based Strategies............................................................................. 172
  8.3.1 Richard’s Method Based on the Laplacian-of-Gaussian Operator..... 172
  8.3.2 Aarnink’s Practical Clinical Method for Contour Determination...... 173
  8.3.3 Pathak’s Edge-Guided Boundary Delineation................................... 174
  8.3.4 Comments on the Edge Based Methods........................................... 175
8.4 Model Based Strategies........................................................................... 176
8.4.1 A Semi-Automatic Method Using Predefined Template Contours . . 176
8.4.2 Ladak’s Model Based Initialisation . . . . . . . . . . . . . . . . . . 176
8.4.3 Other Model Based Methods . . . . . . . . . . . . . . . . . . . . . . 177
8.5 Final Remarks . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 177

9 On Deformable Models 179
9.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 180
9.2 Energy Minimising Formulation . . . . . . . . . . . . . . . . . . . . . . . . 181
9.2.1 The Internal Energy . . . . . . . . . . . . . . . . . . . . . . . . . . . . 181
9.2.2 The External Energy . . . . . . . . . . . . . . . . . . . . . . . . . . . . 183
9.3 Euler’s Equation for Snakes . . . . . . . . . . . . . . . . . . . . . . . . . . . 186
9.4 Numerical Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 187
9.5 Dynamic Deformable Models . . . . . . . . . . . . . . . . . . . . . . . . . . 189
9.5.1 The Lagrangian Formulation of Dynamic Deformable Models . . . 189
9.5.2 Numerical Simulation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 191
9.6 External Forces . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 192
9.6.1 Interactive Constraint Forces . . . . . . . . . . . . . . . . . . . . . . . . 193
9.6.2 Inflation Force . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 193
9.6.3 Distant Attraction Potential Force . . . . . . . . . . . . . . . . . . . . . 194
9.7 Variation and Extension of the Original Snake Model . . . . . . . . . . . 195
9.7.1 Probabilistic Deformable Models . . . . . . . . . . . . . . . . . . . . . 196
9.7.2 Fourier Deformable Models . . . . . . . . . . . . . . . . . . . . . . . . 197
9.7.3 Gradient Vector Flow (GVF) . . . . . . . . . . . . . . . . . . . . . . . . 201
9.8 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 204

10 Contour Extraction Using Multiresolution Snakes 207
10.1 The Motivation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 208
10.2 Multi-Scale Edge Detection . . . . . . . . . . . . . . . . . . . . . . . . . . . 211
10.3 Snake Initialisation: The Importance . . . . . . . . . . . . . . . . . . . . . 214
10.3.1 Traditional Snake . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 214
10.3.2 Balloon . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 215
10.3.3 GVF Snake . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 217
10.3.4 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 217
10.4 Snake Initialisation: A Semi-Automatic Centre-Based Approach . . . . 219
10.4.1 The Radial Profiles . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 220
10.4.2 The Adaptive Neighbourhood . . . . . . . . . . . . . . . . . . . . . . 220
10.4.3 Locating the Seed Points . . . . . . . . . . . . . . . . . . . . . . . . . . 224
10.5 Multi-Scale Snakes: Extracting the Prostate Contour . . . . . . . . . . . 226
10.6 Discussion . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 230
10.6.1 Decomposition Scale of the Dyadic Wavelet Transform . . . . . . . 230
10.6.2 Control Parameters of the Snake . . . . . . . . . . . . . . . . . . . . . . 231
10.6.3 The Anchor Point ............................................. 231
10.6.4 The Seed Points ............................................. 231
10.7 Final Remarks .................................................. 232

11 Quantitative Evaluation of the Contour Extraction Scheme 235
11.1 Pixel Size Calibration ........................................... 236
11.2 Distance-based Metrics ......................................... 236
    11.2.1 The Hausdorff Distance ................................. 236
    11.2.2 Intra-Observer Variability ............................. 238
    11.2.3 Inter-Observer Variability ............................. 244
11.3 Area-based Metrics ............................................ 246
    11.3.1 The “Average” Contour ................................. 247
    11.3.2 Definition of the Accuracy $M_a$ and Sensitivity $M_s$ ........................................... 249
    11.3.3 Assessment of the Accuracy $M_a$ and Sensitivity $M_s$ of CES ............................ 251
11.4 Conclusion .................................................... 253

12 Conclusion 257
12.1 Contributions of Thesis ........................................ 257
12.2 Implications and Suggestions for Future Work .............. 258

A Euler’s Equation for Snakes 261
A.1 Functionals ...................................................... 261
A.2 Fundamental Lemma of the Calculus of Variations ............ 261
A.3 The Euler Equations ............................................ 262
A.4 Euler’s Equations for Variational Problems with Higher Derivatives ........................................... 264
A.5 Euler’s Equation for Snakes .................................... 264

B Lagrange’s Equations of Motion for Dynamic Deformable Models 267
B.1 Euler’s Equations for Variational Problems with Several Unknown Functions 267
B.2 Hamilton’s Principle and Lagrange’s Equations of Motion ........................................... 268
B.3 Euler’s Equations for Variational Problems with Several Independent Variables ........................................... 268
B.4 Lagrange’s Equations of Motion for Dynamic Deformable Models ........................................... 269

Bibliography 271

List of Publications 301
List of Figures

1.1 Key elements of computerised analysis of medical images .............. 2
1.2 Core structure of thesis ............................................. 4
2.1 Diagram of the male pelvis .......................................... 10
2.2 Clinical staging of prostate cancer ................................... 12
2.3 Free and bound prostate specific antigen (PSA) in the bloodstream .... 13
2.4 Biopsy using ultrasound guidance ..................................... 14
2.5 Brachytherapy with ultrasound guidance ............................. 16
2.6 A block diagram of a B-mode ultrasound imaging system .............. 18
2.7 The Proscan Plus imaging system ..................................... 21
2.8 Ultrasound images of a normal prostate gland and that with cancer ... 23
2.9 A sample transverse image of the prostate ............................. 24
2.10 A sample longitudinal image of the prostate ............................ 26
2.11 Split screen showing transverse and longitudinal prostate images ..... 26
2.12 Two more sample ultrasound images of the prostate ................. 27
3.1 The Rayleigh distribution for different parameter values, .............. 31
3.2 The Rician distribution of various parameters. ....................... 32
3.3 The K distribution of various parameters ............................ 34
3.4 The Fisher-Tippett distribution of various parameters ................. 35
3.5 The Nakagami distribution of various parameters ..................... 37
3.6 A plot of the function $f(\beta)$ of relevance to the Weibull distribution. 39
3.7 The Weibull distribution of various parameters ....................... 39
3.8 The generalised Gamma distribution of various parameters ........... 40
3.9 The generalised Fisher-Tippett distribution of various parameters .... 42
4.1 A schematic setup of a hypothetical B-mode ultrasonic imaging system. 54
4.2 Simulated B-scan image of a homogeneous area by the hypothetical B-mode ultrasound imaging system ............................. 58
4.3 Comparison of a horizontal and a vertical line profile .................. 58
4.4 Comparison of the histogram of the simulated image with the fitted p.d.f. of the Rayleigh distribution .............................. 60
4.5 Results of the chi-squared goodness-of-fit test on 500 simulated B-scan images | 62
4.6 Results of the Kolmogorov-Smirnov goodness-of-fit test on 500 simulated B-scan images

Comparison of the boxplots of the goodness-of-fit test statistics of the method of MLE, CDF and PDF. | 64
4.8 Effect of scatterer density: histogram of a simulated image with scatterer density reduced to $d_s = 0.5$. | 66
4.9 A simulated image with “standard” logarithmic compression. | 68
4.10 Effect of logarithmic compression: histogram of a simulated image with ad hoc logarithmic compression applied. | 69
4.11 Results of the chi-squared test on 500 simulated B-scan images with the ad hoc logarithmic compression | 71
4.12 Results of the Kolmogorov-Smirnov test on 500 simulated B-scan images with the ad hoc logarithmic compression | 71
4.13 Simulated B-scan image of a phantom with a square specular component | 73
4.14 Boxplots of the statistics of the goodness-of-fit tests | 75
4.15 Histogram of one of the simulated images of a phantom which contains a square specular component, with ad hoc logarithmic compression applied. | 76
4.16 Boxplots of the statistics of the goodness-of-fit tests | 78
4.17 Simulated B-scan image of a phantom with a small square specular component | 79
4.18 Simulated B-scan image with three specular components | 82
4.19 Histogram of image “us1” | 83
4.20 The histogram of image “us1” and the fitted Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions | 84
4.21 The histogram of image “us6a” and the fitted Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions | 86
4.22 The histogram of image “us6b” and the fitted Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions | 86
4.23 Boxplots of the statistics of the goodness-of-fit tests | 87

5.1 Illustration of the concepts of hard and soft thresholding, and linear and nonlinear shrinkage of wavelet coefficients. | 104

6.1 Illustration of the concepts of multiresolution analysis | 118
6.2 Illustration of the filtering scheme for a two-stage multiresolution analysis | 122
6.3 Illustration of the relationship between the subspaces of a biorthogonal wavelets system | 124
6.4 Waveforms of the quadratic spline scaling function and wavelet | 126
6.5 A noise corrupted signal and its DWT | 129
6.6 A simulated B-scan ultrasound image and its approximation | 130
6.7 The DWT details of the simulated image | 131
LIST OF FIGURES

6.8 The modulus of the dyadic wavelet transform of the simulated image ... 133
6.9 Modelling the distribution of the wavelet coefficients as a two-state zero-mean Gaussian mixture ... 136
6.10 Modelling the distribution of the wavelet coefficients as a two-state zero-mean Gaussian mixture ... 138
6.11 Modelling the distribution of the modulus as a two-state Rayleigh mixture ... 140
6.12 Modelling the distribution of the squared modulus as a two-state Gamma mixture ... 141
6.13 Illustration of \(a\ posteriori\) probability function \(m_j\) ... 143
6.14 Illustration of wavelet coefficient shrinkage function \(\hat{m}_j\) ... 145
6.15 The noise-reduced image and its difference with the original image ... 147
6.16 Comparison of the histograms of the original image and the noise-reduced image ... 147
6.17 Simulated image denoised by median filters ... 148
6.18 Simulated image denoised by Wiener filters ... 149
6.19 Simulated image denoised by Lee filters ... 149
6.20 Comparison of original ultrasound image “us1” and after denoised by the noise reduction scheme ... 152
6.21 Difference between the original ultrasound image “us1” and that denoised by the noise reduction scheme ... 153
6.22 Another example of application of the noise reduction scheme ... 154
6.23 Difference between the original ultrasound image “us5” and that denoised by the noise reduction scheme ... 155
7.1 Illustration of a LoG operator ... 158
7.2 Binary thresholded images after applying LoG operators of varying \(\sigma\) ... 160
7.3 Result of applying the LoG operator to the image “us1” ... 162
7.4 Result of applying the LoG operator to the image “vs04” ... 164
9.1 Effects of snake regularisation parameters \(\alpha\) and \(\beta\) ... 183
9.2 The influence of the external energy ... 185
9.3 A contour decomposed into component ellipses ... 198
9.4 A “mean” contour and its two variations ... 199
10.1 A schematic overview of the proposed contour extraction scheme ... 208
10.2 The external force plays the important role ... 210
10.3 The dyadic wavelet transform of the prostate ultrasound image “us1” ... 213
10.4 A snake initialised not sufficiently close to the desired boundaries has poor convergence ... 216
10.5 The external force field derived from the modulus of the wavelet transform ... 217
10.6 The performance of a balloon is no better than that of a traditional snake ... 218
10.7 The performance of GVF snakes is not encouraging either ... 219
10.8 The centre based approach: extracting the radial profiles . . . . . . . . . . . 221
10.9 Gradients and edge vectors are perpendicular to one another. . . . . . . . . 222
10.10 The edge vectors of image “us1” at scale $2^4$ . . . . . . . . . . . . . . . . . 223
10.11 Locating the seed points by searching for local maxima in radial profiles . . 224
10.12 Typical seed points identified at scale $2^4$. . . . . . . . . . . . . . . . . . . 225
10.13 Multi-scale snake deforming and extracting the contour of the prostate . . 227
10.14 Final contour extracted for image “us1” . . . . . . . . . . . . . . . . . . . . . 228
10.15 Results of applying the proposed contour extraction scheme to image “vs01”. 229
10.16 Final contour extracted for image “vs01” . . . . . . . . . . . . . . . . . . . 230

11.1 Intra-observer variability: comparison of the boxplots of the Hausdorff dis-
tance $D_H$ of MO1, MO2 and CES . . . . . . . . . . . . . . . . . . . . . . . . . . . 239
11.2 Comparison of the empirical cumulative distribution functions of the Haus-
dorff distance $D_H$ of MO1, MO2 and CES . . . . . . . . . . . . . . . . . . . . . 240
11.3 Comparison of the boxplots of the sample means $D_M$ . . . . . . . . . . . . . 243
11.4 Inter-observer variability: comparison of the boxplots of the Hausdorff dis-
tance $D_H$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 245
11.5 Comparison of the boxplots of the sample means $D_M$ . . . . . . . . . . . . 247
11.6 The “average” contours of the three sets of segmentation results of image
“us1” . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 248
11.7 Definition of the different regions enclosed by the “average” contour of the
manual outlining and the contour extracted by the CES . . . . . . . . . . . . . . . 250
11.8 Comparison of the “average” contour of manual outline and the contour
extracted by the CES for image “us1”. . . . . . . . . . . . . . . . . . . . . . . . 253
11.9 Comparison of the boxplots of the sensitivity $M_s$ of CES with that of the
two human operators . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 254
11.10 Comparison of the boxplots of the accuracy $M_a$ of CES with that of the
two human operators . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 255
List of Tables

2.1 Gleason grading of cancer ........................................ 15
2.2 Specifications of the Proscan Plus imaging system ............. 22

4.1 Results of the Wilcoxon-Mann-Whitney test on the chi-squared goodness-of-fit test statistics ........................................ 64
4.2 Results of the Wilcoxon-Mann-Whitney test on the K-S goodness-of-fit test statistics ........................................ 64
4.3 Number of rejections of the null hypothesis that simulated B-scan images follow the K distribution ........................................ 67
4.4 Number of rejections of the null hypothesis that the simulated B-scan images in Case 1 follow the K distribution or the generalised Gamma distribution 74
4.5 Number of rejections of the null hypothesis that the simulated B-scan images in Case 2 follow the specified distributions .......................... 76
4.6 Number of rejections of the null hypothesis that the simulated B-scan images in Case 3 follow the specified distributions .......................... 77
4.7 Number of rejections of the null hypothesis that the simulated B-scan images in Case 4 follow the specified distributions .......................... 80
4.8 Number of rejections of the null hypothesis that the simulated B-scan images in Case 5 follow the specified distributions .......................... 81

6.1 Coefficients of the FIR filters associated with the quadratic spline wavelets and the scaling functions ........................................ 127
6.2 Comparison of the SNR of the original simulated ultrasound images and the denoised images ........................................ 151
6.3 Comparison of the SNR of the original ultrasound images of the prostate and the denoised images ........................................ 155

8.1 Summary of cited references on the problem of prostate contour extraction 168

10.1 Parameters of a traditional snake, a balloon and a GVF snake ........... 215
10.2 Parameters of a multi-scale snake ....................................... 226

11.1 Intra-observer variability: comparison of the Hausdorff distance $D_H$ .... 238
11.2 Results of the Kruskal-Wallis test on the Hausdorff distance $D_H$ . . . . . . 241
11.3 Multiple comparison of difference of mean ranks (M.R.D.) of the Hausdorff distance $D_H$ . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 242
11.4 Comparison of the statistics of the sample means $\overline{D}_M$ . . . . . . . . . . 243
11.5 Comparison of the differences of the means (M.D.) of the sample means $\overline{D}_M$ 243
11.6 Inter-observer variability: comparison of the Hausdorff distance $D_H$ . . . . . 245
11.7 Comparison of medians and means of the sample means $\overline{D}_M$ . . . . . . . . 246
11.8 Assessment of CES by area-based metrics: sensitivity $M_s$ . . . . . . . . . . 251
11.9 Assessment of CES by area-based metrics: accuracy $M_a$ . . . . . . . . . . . 252
I am most grateful to my supervisors Dr. Chris deSilva and Professor Yianni Attikiouzel for their tangible support, constant encouragement and guidance, without which this thesis would never have come to be.

To Dr. Chris deSilva I am deeply grateful for his penetrating mathematical insight and precise guidance whenever I needed throughout this study, leaving me enough space to explore, but never letting me get lost, just like a light house glowing in the dark, which is vital to a vessel navigating in the sea. In particular I am deeply grateful for his meticulous proofreading and criticism of early draft of this thesis. I will never forget the days he generously spent with me in front of the computer examining the images. Many thanks, Chris.

To Professor Yianni Attikiouzel I am grateful for his unambiguous advice and constant encouragement. I still remember the afternoon when he showed me how to write a PhD thesis on the white-board in his office, from the viewpoint of a master at this art. Many thanks, Yianni.

To my fellow postgraduate students at the Centre for Intelligent Information Processing Systems (CIIPS), Frederick Howe-Hui Chee, Thomas Hanselmann, Pantazis Houlis, Jordan Kosek, Martin Masek, Pubudu N. Pathirana, and James Hao-Pin Young, I am thankful for their friendship, assistance and humor. Cheers, guys.

I am also grateful to CIIPS administrative staff, Ms Violetta Cetrullo, Mrs Brenda Churchill, and Ms Sandra Passamani for their various support throughout my candidature. Thanks also go to Dr. Ramachandran Chandrasekhar for his helpful comments and discussion, and especially the \LaTeX{} templates files, which have made the thesis preparation much easier. I would like to thank Dr. Hugh Dawkins of Department of Urology, UWA, for providing me with the ultrasound images used in this study and various support. I am also thankful to Chenyang Xu and J.L. Prince of the Image Analysis and Communications Lab, The Johns Hopkins University, whom I have not met before, for their implementation of the gradient vector flow (GVF) snake (downloadable from http://iacl.ece.jhu.edu/projects/gvf/) used in Chapter 10.

Thanks also go to Associate Professor Weiping Hu of Guangxi Normal University, for exchange of ideas, laughter and his friendship during his visit to CIIPS, and Associate Professor Xin Xu of Wuhan University (visiting CIIPS now) for bringing to my attention some references on SAR speckle reduction and his suggestion on organisation of the thesis.
It is my pleasure to acknowledge the financial support of an Overseas Postgraduate Research Scholarship (OPRS) provided by the Australian Commonwealth Government, and additional financial support provided by The University of Western Australia (UWA) in the form of University Postgraduate Award (UPA), which together made this study possible. I am also grateful to CIIPS and the School of Electrical, Electronic and Computer Engineering for granting me the financial assistance and opportunity to attend relevant conferences. I also thank the School for giving me the opportunity to work as a teaching assistant during the second half of the year 2001. I would also like to take this opportunity to thank all participants of CIIPS Postgraduate Seminars.

To Professor Kenneth Clements (my wife Yihui Lan’s PhD supervisor) and his wife Professor H.Y. Izan I am grateful for their kind support to our family during the past few years. I am also grateful to my friends over the years, Dr. Dehong Yu, Ms Huayou Wu, Associate Professor Caiheng Li, Ms Jianping Wu, Dr. Wanquan Liu, Ms Wen Wang, Dr. Wangqing Li, Ms Mei Liu, Dr. Feng Gao, Dr. Wenge Liu, Ms Ying Guo, Dr. Sheng Chang, Ms Wei Zong, Dr. Chunnong Zhao, Dr. Jian Li, Mr. Hongyin Fan, Ms Hongmei Sun for their sincere friendship and kind assistance to my family (there are numerous people who are not listed here, but in my heart).

Finally, I owe my parents and parents-in-law a lot, who sacrificed their otherwise independent and tranquil lives after retirement to bear the difficulties of coming to Perth to help me and my wife, who otherwise could not survive the dual pressures of pursuing PhD degrees and family life. I thank my parents for their love and care and especially for instilling in me the desire, confidence, and perseverance to pursue my dreams. My brother and sister I thank them for their love and encouragement over the years. My beloved wife Yihui, who has been putting up with many bad moods and excuses that I was “too busy” over the years, and our two wonderful boys Han-Qing and Yuan-Long, who have been “enjoying” daddy-free evenings most of the time, I owe a special debt which I will pay back in my whole life.

Fangwei Zhao
Perth, Western Australia
November, 2003
Chapter 1

Introduction

Cancer remains a major public health problem in the world. One in four deaths in the United States is caused by cancer. In recent years, prostate cancer has become a new focus of male health due to the aging and expanding population. In the United States prostate cancer is one of the most commonly diagnosed cancers and a leading cause of death of older men, second only to lung and bronchus cancer. It is estimated more than 220,900 new prostate cancer cases will be diagnosed in the United States in the year 2003, which accounts for nearly 33% of new cancer cases in men [1].

Since prostate cancer has posed a new threat to the health of men, many efforts have been made to lower the mortality rate from this disease. However, in the absence of a cure for advanced prostate cancer, the most feasible approach may be to improve the detection rate of the disease in early diagnosis. When it is diagnosed in an early stage, prostatic carcinoma is curable [2].

A variety of imaging tools are available for diagnosing and staging prostate cancer, including computed tomography (CT), magnetic resonance imaging (MRI) and ultrasound. Until recently, CT was the most accurate. However, it is subject to significant limitations as prostate cancer cannot be directly seen by noncontrast or contrast-enhanced CT [3]. MRI is unique in its ability to stage prostate cancer and identify the surrounding periprostatic structures, and hence has the greatest potential to accurately assess tumor extension [3]. Ultrasound can identify tumor within the prostate. Despite its inability to consistently identify subtle tumor beyond the prostate, it remains an economical and valuable tool in clinical practice.

Although prostate ultrasound, particularly the transrectal ultrasound (TRUS) technique, was developed over 25 years ago, it has found wide application only in the last decade. With the development of high definition ultrasound probes, TRUS has become one of the most valuable tools in early detection and management of prostatic cancer thanks to its real-time capability, low cost, portability, ease of use, minimal discomfort, and non-invasive nature [4–6]. However, the interpretation of a prostate ultrasound image depends critically on the experience and expertise of a urologist (let alone the effects of tiredness, boredom and other human variability) and is still difficult and subjective. “Al-
Figure 1.1: Key elements of computerised analysis of medical images (adapted from [10]).

though a great deal of knowledge” on TRUS “has been gathered over the last few years, there is still much to be learned”, as noted by an expert urologist [3].

To overcome the subjective interpretation and facilitate objective diagnosis, computer aided analysis of ultrasound images of the prostate would be very helpful. Computer aided analysis of images may improve diagnostic accuracy by providing a more reproducible interpretation of the images including textural information that human observer may find difficult to perceive (the dynamic range of the human vision system is only around 30dB).

1.1 Computerised Analysis of Medical Images

Computer aided analysis of medical images, especially computer aided diagnosis, has drawn much attention during the last few decades [7], as evidenced by the special issue of the IEEE Transactions on Medical Imaging [8] published recently. Although X-ray, ultrasound, and magnetic resonance imaging (MRI) remain the main medical image modalities that computerised image analysis has been involved with, less well-known ones such as thermography and Moire imaging are also beginning to benefit. With the advances in computer aided diagnosis which are enthusiastically welcomed by both the medical profession and the public, “in the future, it is quite likely that all medical images will undergo some form of computer analysis in order to benefit the diagnosis”, as predicted optimistically by the guest editors of the special issue [9].

As illustrated in Figure 1.1, computerised analysis of medical images in general may include the following key elements:

Medical imaging modalities, which may at least include X-ray computed tomography (CT), MRI, ultrasound, positron emission tomography (PET), and emission computed tomography (ECT). Most of the time, the details of these medical imaging technologies are ignored since most researchers regard them as black-boxes or data feeders, and are merely concerned about the outputs, namely, the medical images they generate. However, it is beneficial (sometimes essential) to understand the underlying physical principles and image formation processes behind these imaging technologies,
especially for the subsequent procedures such as preprocessing.

Preprocessing, which aims to remove or suppress irrelevant information and/or enhance the image to improve visualisation or facilitate processing at later stages. Examples may include histogram equalisation for contrast enhancement or the nontrivial task of noise suppression. Improvements at this stage may substantially influence performance of the procedures at later stages such as segmentation, and hence the overall performance.

Feature extraction and segmentation, which is intended to extract regions of interest or divide an image into mutually exclusively regions for the purpose of tissue characterisation or object classification later. This may be accomplished in the original spatial domain or in a feature space which involves feature nomination and extraction, or other mathematical transformation through which pixels of different regions can be decorrelated or discriminated. A tangible example may be segmenting an ultrasound image of the prostate into prostate region and non-prostate region. This is one of the most important and difficult procedures.

Classification, characterisation, and identification, which establish the physical link between each region segmented and the underlying anatomical or pathological structures. This procedure depends critically on a priori knowledge about the anatomical or pathological structures of concern and their corresponding presentation in the medical images.

2D/3D visualisation, which may follow any of the above procedures. It is aimed at providing a radiologist with information that is not readily available or hard to perceive in the original images. This may be the output of preprocessing such as an enhanced image with noise suppressed, or the result of segmentation such as possible lesion regions highlighted with pseudo colour, or a 3D presentation of the object of interest, which may facilitate the subsequent quantitative and/or qualitative analysis, and ultimate diagnostic decisions.

Quantitative and/or qualitative analysis, which may involve measurement of areas and/or volumes of anatomical or pathological structures, which is quantitative analysis. Sometimes qualitative analysis depends on quantitative analysis. A possible scenario of this is to infer whether the prostate is normal or abnormal on the basis of its volume and symmetry, which are quantifiable.

Diagnosis report, which may in the form of diagnostic decision or some annotated images which have treatment implications such as surgery planning.

Knowledge base, which prevails in the whole spectrum of computer aided analysis of medical images. For example, a priori knowledge about the statistical properties of noise proves important in developing noise suppression techniques. It can also take the
form of anatomical and pathological information about tissues, which may facilitate final diagnostic decision.

By convention, the above procedures up to the stage of feature extraction and segmentation are known as low-to-medium level processing, as opposed to higher level procedures such as classification and identification, which are more “human”. They are essential to the subsequent procedures of computer aided analysis of medical images, and will be the focus of this thesis.

1.2 Structure of Thesis

This thesis is an attempt to address the first few key elements of computer aided analysis of B-mode ultrasound images of the prostate. Specifically, we endeavour to

1. model formation of B-mode ultrasound images and their statistical properties;
2. suppress ultrasound speckle; and
3. extract the contour of the prostate.
INTRODUCTION

These three tasks respectively correspond to issues about medical imaging modality, pre-processing, and segmentation in computer aided analysis of medical imagery in general (see Figure 1.1).

This thesis consists of twelve chapters and two appendices. Figure 1.2 illustrates the core structure of the thesis excluding Chapter 1 (Introduction), Chapter 2 (Ultrasound of the Prostate: The Background) and Chapter 12 (Conclusion).

Coherently these chapters respectively elaborate the three tasks listed above, namely, Chapters 3 and 4 for ultrasound image modelling, Chapters 5 and 6 for ultrasound speckle reduction, and Chapters 7-11 for prostate contour extraction.

Functionally these chapters fall into four categories in sequence (as indicated by the four columns divided by vertical lines in Figure 1.2), respectively entitled literature review, mathematical background, methodologies and experiments, and evaluation (of proposed methods). For each task we elaborate how it was dealt with by following this functional “road-map”, namely, we first review the need or motivation of the task and the current literature, then lay the mathematical foundation for our methodologies and experiments, which are elaborated later and finally evaluated.

Some chapters serve a single purpose, such as Chapter 5 in which we review the current literature on speckle reduction. Some chapters are “multi-purpose”. For example, Chapter 6 is a “three-in-one” chapter, introducing the mathematical background on wavelets in section 6.2, elaborating a speckle reduction scheme using the dyadic wavelet transform in section 6.3, and evaluating the speckle reduction scheme in section 6.4.

1.3 Outline of Thesis

We focussed on the structure of the thesis in the last section, where we clarified the functional and logical connections of the chapters in an attempt to familiarise the reader with the thesis as a whole “picture”. We now outline the thesis below (refer back to Figure 1.1 and 1.2):

Chapter 2 is a concise review of the scientific and medical background of the prostate and ultrasound imaging of the prostate. Section 2.1 first presents the necessary medical background to the prostate gland, prostatic disease, especially prostate cancer, and then focuses on the role ultrasound plays in diagnosis and treatment of prostate cancer. In section 2.2 we introduce B-mode ultrasound imaging with emphasis on its signal-processing aspects. The ultrasound images of the prostate used in this study were collected from a commercial Proscan Plus prostate imaging system, which will be described in section 2.3. In section 2.4 we attempt to explain the appearance of ultrasound images of the prostate, specifically, what tissues are or what makes tissues hypoechoic or hyperechoic. Finally presented in section 2.5 are the sample ultrasound images of the prostate used throughout this thesis.
1.3. OUTLINE OF THESIS

1.3.1 Ultrasound Image Modelling

B-scan ultrasound images usually exhibit low resolution with a high level of speckle noise due to scattering and other complicated interactions between ultrasonic pulses and human tissue. Suppression of speckle noise is desirable in order to improve visualisation or the reliability of later feature extraction and segmentation processes, which are essential to the computer aided analysis of images. Before any attempt is made to deal with speckle, it would be beneficial to gain a better understanding of its nature by examining ultrasonic pulse-echo imaging and consequently analysing the statistical properties of speckled images.

In Chapter 3 we review some distribution models of speckle, which has been treated extensively in the literature for laser, synthetic aperture radar (SAR) and ultrasound imagery. We also empirically propose the generalised Fisher-Tippett distribution, which will be used to analyse the ultrasound images of the prostate in section 4.3. We introduce in section 3.2 methods for parameter estimation of these distribution models, including the method of maximum likelihood estimation (MLE), fitting the probability density function (PDF), fitting the cumulative distribution function (CDF) and the method of moments. These methods allow us to analyse the statistical properties of simulated B-scan images by fitting their histograms to different distribution models, as described in Chapter 4. Section 3.3 presents the goodness-of-fit test procedures, including the chi-squared test and the Kolmogorov-Smirnov test, which will be used in Chapter 4 to test our hypotheses and verify whether the pixel values of simulated images really follow various specified distribution models for different scenarios.

In Chapter 4 we present a computer simulation of B-mode ultrasound imaging, which provides us not only with an insight into the nature of speckle, but also with a viable test-bed for our speckle reduction scheme later. Several questions will be addressed by the computer simulation, such as how the B-scan images are formed from the backscattered pulse-echo signal, and whether the speckle is solely a characteristic of the imaging system or a result of interaction between the tissue and the imaging system. In section 4.1 a hypothetical B-mode ultrasonic imaging system will be described, the features of which include transducer modelling and tissue modelling, envelope detection, logarithmic compression and simulation of scatterer density. Section 4.2 presents simulated B-scan images for various scenarios and the analysis of their statistical properties, which inspired the statistical analysis of the ultrasound images of the prostate, described in section 4.3.

1.3.2 Ultrasound Speckle Reduction

Speckle refers to the granular appearance that compromises the image quality and resolution in optics, synthetic aperture radar (SAR), and ultrasound. Extensive efforts have been devoted to counteracting the adversity of speckle by the optical, SAR, and ultrasound imaging community. In Chapter 5 we review these attempts, which generally fall into two categories, namely, those at the system level (pre-image-formation) and those at
INTRODUCTION

the image level (post-image-formation). We first review in section 5.1 system level speckle reduction techniques such as compound scanning, and then focus on post-image-formation processing techniques in the remaining text, which we loosely classify into classic adaptive filters, wavelet methods, geometry driven diffusion and other methods.

In Chapter 6 we present a speckle reduction scheme based on Mallat and Zhong’s dyadic wavelet transform (MZDWT) and statistical modelling of the wavelet coefficients. A brief introduction to wavelets and multiresolution analysis in general, and the quadratic spline wavelet and the corresponding dyadic discrete wavelet transform in particular is given in section 6.2. The proposed speckle reduction scheme is then described in section 6.3, which in on the basis of modelling statistical properties of the wavelet coefficients and exploiting their inter-scale correlation.

Specifically, the squared modulus of the component wavelet coefficients are modelled as a two-state Gamma mixture. Inter-scale correlation is exploited by taking the harmonic mean of the posterior probability functions, which are derived from the Gamma mixture. Results of applying the noise reduction scheme to simulated B-scan ultrasound images and real ultrasound images of the prostate are presented in section 6.4. Its performance is quite satisfactory. The important features of the original images are preserved while most of the speckle noise is removed. It is also evaluated both qualitatively and quantitatively in comparison with median, Wiener, and Lee filters, and its performance surpasses that of all these filters.

1.3.3 Prostate Contour Extraction

It is essential to accurately extract prostate contours in many diagnostic and treatment procedures for prostate disease. However, due to low resolution, poor contrast and high noise levels, it is a difficult task to locate the contour of the prostate in ultrasound images automatically by computer. As a pilot study, we approach this problem in Chapter 7 by applying the Laplacian of Gaussian (LoG) operator to detect the contour of the prostate. This is a valuable experience which motivates us to take a multi-scale perspective later in Chapter 10 to design a contour extraction scheme in the framework of multiresolution analysis.

Before that, we review in Chapter 8 some solutions in the literature to the problem of delineating prostate contours in B-scan ultrasound images. Based on the underpinning strategies, they can be roughly classified into three groups: pixel or texture based (section 8.2), edge based (section 8.3), and model based (section 8.4). We also summarise the main features and disadvantages of these solutions.

Deformable models are a set of promising and actively researched model-based image analysis techniques which combine physics, geometry and approximation theory. Two-dimensional deformable models (snakes) have found wide applications in a variety of medical image analysis problems, such as segmentation, visualisation, and tracking of biological structures. In Chapter 9 we introduce the mathematical background of deformable
1.3. OUTLINE OF THESIS

models and outline their general applicability to medical image interpretation to lay the foundations for the contour extraction scheme to be described in Chapter 10. For the sake of completeness, we briefly present Euler’s equation for snakes in Appendix A, and introduce the background of Hamilton’s principle and Lagrange’s equations of motion in Appendix B.

Based on multiresolution analysis (MRA), in Chapter 10 we propose a novel contour extraction scheme (CES), which fuses MZDWT and snakes. It is inspired by the previous work done by other researchers in this area and especially the pilot study in Chapter 6 and 7, which (despite the different goals) motivates us to put the problem of extracting the contour of the prostate in a multi-scale framework provided by the wavelet transform.

Traditionally, the external potential function of a snake is defined only once and for all at a fixed resolution. The snake evolves only within this specific external force field, which once decided on, cannot be changed during the deformation of the snake. In contrast, we propose a multi-scale snake, whose external potential functions are designated as the edge maps at different resolutions provided by the MZDWT, and thus are “switchable”. The deformation and migration of the snake from coarse to fine scales, gradually fending off the interference from noise, eventually extracts the contour of the prostate.

It is important to initialise a snake fairly close to the desired features. Accordingly the snake is initialised using some “seed points”, which are identified semi-automatically. To keep the interference of the noise to a minimum, seed point identification begins with the edge map at the coarsest scale. An anchor point is selected manually near the centre of the prostate region and the seed points are then identified automatically as the maxima along the radial profiles which emanate from the anchor point. A snake initialised using these seed points evolves across the edge maps at different scales and finally converges to the contour of the prostate.

Although our CES successfully defines the prostate contours which closely match what a human observer would infer from the original images, it is necessary to evaluate it quantitatively to ensure that its performance is reliable, and no worse than that of the manual outlining process. In Chapter 11 we first use distance-based and then area-based criteria to assess the segmentation results of our CES against the boundaries manually outlined by human operators. The distance-based metrics are defined on the basis of the Hausdorff distance and are used to assess the reproducibility or consistency of the CES. The area-based metrics are used to evaluate its accuracy and sensitivity, which depends on the area enclosed by the contour extracted.

Finally Chapter 12 concludes the thesis with discussions, implications, and suggestions for future work. It is interesting to note that we take a multiresolution viewpoint throughout the thesis, either in dealing with the task of speckle reduction or with segmentation. Hence the keyword in the title of this thesis.
Chapter 2

Ultrasound of the Prostate: The Background

In this chapter we review the scientific and medical background on the prostate and ultrasound imaging of the prostate. In section 2.1 we present the medical background\(^1\) to the prostate gland, prostatic disease, especially prostate cancer, and the role ultrasound plays in diagnosis and treatment of prostate cancer. Section 2.2 is a general review of the signal processing aspects of B-mode ultrasound imaging. In section 2.3 we introduce the commercial Proscan Plus prostate imaging system, from which the ultrasound images of the prostate used in this study were collected. Section 2.4 is an attempt to explain the appearance of ultrasound images of the prostate, namely, what tissues are or what makes tissues hypoechoic or hyperechoic. Finally we present in section 2.5 the sample ultrasound images of the prostate used in this thesis.

2.1 The Prostate

The prostate is a small organ located deep in the male pelvis. It is situated in front of the rectum, just below the bladder and surrounds a portion of the urethra, the tube that carries urine from the bladder out through the penis. Figure 2.1 shows a diagram of the male pelvis illustrating the position of the prostate gland (courtesy of The Prostate Centre, University of Toronto, Canada [11]). The prostate gland is about the size and shape of a walnut, and weighs between 12 and 20 grams in a normal young man [3].

2.1.1 Function of the Prostate

The prostate gland is part of a man’s reproductive system and noted for its exocrine function. Specifically, it is made up of thousands of tiny fluid-producing glands, and the fluid that the prostate secretes makes up semen, the fluid that is stored with sperm in the seminal vesicles and carries sperm during ejaculation. When the male climaxes, muscular

\(^1\)This section is adapted from [11], courtesy of The Prostate Centre, University of Toronto, Canada.
2.1. THE PROSTATE

contractions also cause the prostate to secrete this fluid into the urethra, where sperm is expelled from the body through the penis.

The prostate also produces an enzyme (a kind of protein) called prostate specific antigen (PSA). PSA is released with the ejaculatory fluid and is also absorbed into the bloodstream. The testing of PSA levels in the blood is used to detect prostate cancer.

Apart from its notable role in producing ejaculate, the prostate also plays a part in controlling the flow of urine. The prostate surrounds the urethra which passes from the bladder to the penis. Contraction of muscular fibres in the prostate can slow the flow of urine [11]. It also blocks the flow of urine at the time of ejaculation.

2.1.2 Prostatic Disease

Although the prostate gland is quite small, it can cause immense clinical problems. One quarter of all men who experience urinary problems actually have prostatitis, which is a condition that causes inflammation of the prostate.

Prostatitis

There are four different types of prostatitis, including acute bacterial prostatitis (ABP), chronic bacterial prostatitis (CBP), non-bacterial prostatitis (NBP), and pelvic floor myalgia (PFM) (or prostatodyinia). Although there are many theories about why a man may contract prostatitis, the causes are not clearly known [11].

Benign Prostatic Hyperplasia (BPH)

The prostate enlarges as men age. For many men, enlargement of the prostate gland is never a problem. All men who live into their 40s and 50s will develop some type of histologic change, however, some of them will undergo pathologic change, and benign
prostatic hyperplasia\(^2\) (BPH) is one of the most common benign processes affecting the gland [3].

The term “hyperplasia” indicates there is an increase in the number of prostate cells including muscle cells, which contribute to an enlargement of the prostate gland. The word “benign” means these cells are not cancerous. In other words, BPH is an enlarged, but otherwise normal prostate. However, as the prostate grows larger, it may narrow the urethra and interfere with normal urine flow. Left untreated, BPH can result in permanent damage to the urinary system even affecting the kidneys. Moreover, although BPH is not cancer, a man can have BPH and prostate cancer at the same time.

**Prostate Cancer**

Our cells normally undergo a predictable cycle of growth and renewal, which is controlled by DNA, the genetic material contained within each cell. When DNA fails to control cells, they may become abnormal and begin to grow and reproduce in an uncontrolled manner. Gradually the body loses its ability to control the behaviour of these abnormal cells, which are described as malignant or cancerous. With time cancerous cells may spread, forming a tumor, or metastasise to other locations in the body [11].

**Clinical Staging of Prostate Cancer**

A common system used by doctors to determine the stage of prostate cancer is the so-called TNM (tumor, node, metastasise) system. According to

T: the type of tumor;

N: tumor spread to the lymph nodes; and

M: tumor spread to distant sites

prostate cancer can be staged as [11]

T1: (also known as Stage A, see Figure 2.2(a)) a small tumor that is confined to the prostate, and not detected during a digital rectal exam (DRE). Prostate cancer at this stage usually produces no symptoms. Treatment may involve observation, surgery to remove the prostate or radiation.

T2: (also known as Stage B, see Figure 2.2(b)) the tumor is confined to the prostate gland but may be detected during a DRE. Possible symptoms may include a need to urinate frequently, especially at night. Treatment may involve surgery or radiation, possibly combined with hormone therapy to shrink the tumor.

T3: (also known as Stage C, see Figure 2.2(c)) the tumor has begun to spread beyond the prostate to areas surrounding the gland. Possible symptoms are similar to those

\(^2\)Or benign prostatic hypertrophy.
of Stage T2. Treatment may involve radiation combined with hormone therapy, or for some patients surgery to remove the prostate.

T4: (also known as Stage D1, see Figure 2.2(d)) the tumor has spread beyond the prostate into surrounding structures including the pelvic lymph nodes (N+). Possible symptoms may include a need to urinate frequently, painful, obstructed urine flow (blood may appear in urine), and fatigue. Treatment may involve hormone therapy, possibly with radiation to ease symptoms.

M+: (also known as Stage D2). When cancer has spread beyond the prostate it is said to have metastasised. Cancer may spread to distant sites such as bone, liver or lung via the lymphatic system and the circulation. Cancer that has reached this stage may produce symptoms such as difficult or painful urination, obstructed urine flow, pain if the cancer has spread to the bone, and fatigue. Treatment for this stage of cancer is aimed at easing symptoms and slowing the progression of the disease. This usually includes hormone therapy.
2.1.3 Diagnosis of Prostate Cancer

Various diagnostic tests have been developed to detect prostate cancer, including the digital rectal examination (DRE), the prostate specific antigen (PSA) test, biopsy under guidance of ultrasound, X-ray CT scan and bone scan.

Digital Rectal Examination (DRE)

The digital rectal examination (DRE) is a common screening procedure for prostate cancer, whereby a physician inserts a gloved, lubricated finger into the rectum, in order to feel the size and shape of the prostate through the rectal wall.

Prostate specific antigen (PSA) Test

Prostate specific antigen (PSA) absorbed into the bloodstream may become bound to two proteins, namely, anti-chymotrypsin (ACT) and alpha macroglobulin (aMG). The PSA test measures the level of free and bound PSA in the blood (see Figure 2.3). If elevated, the antigens may be an indicator of prostate cancer. Combined with the DRE, the PSA test is the most accurate guide to early detection of prostate cancer.

Normally, the level of PSA detected in the blood is between 0.0 and 4.0. However, an increase in the level of PSA does not necessarily mean that one has prostate cancer, just as a low PSA level does not necessarily mean that one does not have prostate cancer. An elevated PSA may be caused by non-cancerous conditions such as BPH, urinary tract infection, prostatitis or any other condition or diagnostic test that might irritate the prostate gland. PSA may even increase slightly after sexual intercourse.

Most men who have been diagnosed with prostate cancer have a PSA level greater than 4.0. Generally, the higher the PSA, the larger the cancer. Although an elevated PSA
2.1. THE PROSTATE

Figure 2.4: Biopsy using ultrasound guidance (Courtesy of The Prostate Centre, University of Toronto, Canada [11]).

does not necessarily mean that one has cancer, it is currently the most accurate indicator that something may be wrong and that further tests such as an ultrasound with biopsy should be performed.

Biopsy

It has become a standard clinical practice to perform a biopsy using ultrasound guidance. A transrectal ultrasound (TRUS) is performed by placing a lubricated probe in the rectum, which produces a B-mode ultrasound image showing the prostate and its parts, including areas where there is cancer. In a prostate biopsy, tiny pieces of tissue are removed from the prostate using a small spring-loaded needle. The needle unit is attached to the ultrasound probe and enters the prostate through the rectum (see Figure 2.4). Usually about 6-12 tissue samples are obtained and then examined under a microscope by a pathologist to see if cancer is present.

Gleason Grade

The tissue samples taken from a patient’s prostate during biopsy are examined under a microscope by a pathologist. A grade of one (low grade) to five (high grade) are assigned to the two most common patterns of cancer seen under the microscope:

1. How the cells look (on a scale of 1 to 5); and

2. How the cells are arranged (on a scale of 1 to 5).

These two numbers are then combined to give a Gleason Grade score of 2-10. Normal prostate tissue consists of a structured arrangement of small glands (acini) and ducts. Low grade cancer is the least aggressive and most resembles normal tissue. As the cancer cells multiply and spread both the appearance and arrangement of the cells will change. The cells change shape and begin to look less like typical prostate cells. This increase in
This is the least dangerous type of cancer. The cells look most like normal prostate cells and are described as being “well-differentiated”. This type of cancer tends to be slow growing.

Intermediate cancers may behave like low-grade or high-grade cancers. The cells’ behaviour may depend on the volume of the cancer and the PSA level. This is the most common grade of prostate cancer.

High-grade cancers are usually very aggressive and quick to spread to the tissue surrounding the prostate. These cancer cells look least like normal prostate cells and are usually described as “poorly-differentiated”.

<table>
<thead>
<tr>
<th>GRADE</th>
<th>SCORE</th>
<th>GROWTH</th>
<th>COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>2–4</td>
<td>Slow</td>
<td>This is the least dangerous type of cancer. The cells look most like normal prostate cells and are described as being “well-differentiated”. This type of cancer tends to be slow growing.</td>
</tr>
<tr>
<td>Medium</td>
<td>5–7</td>
<td>Unpredictable</td>
<td>Intermediate cancers may behave like low-grade or high-grade cancers. The cells’ behaviour may depend on the volume of the cancer and the PSA level. This is the most common grade of prostate cancer.</td>
</tr>
<tr>
<td>High</td>
<td>8–10</td>
<td>Aggressive</td>
<td>High-grade cancers are usually very aggressive and quick to spread to the tissue surrounding the prostate. These cancer cells look least like normal prostate cells and are usually described as “poorly-differentiated”.</td>
</tr>
</tbody>
</table>

Table 2.1: Gleason grading of cancer (adapted from [11], courtesy of The Prostate Centre, University of Toronto, Canada).

All of these changes also cause the ducts and acini of the prostate to take on an irregular appearance. The cells of high grade cancer tend to behave very aggressively. Consequently, normal cells die off as they compete with the cancer cells for nutrients. As high grade cancer cells continue to spread, the ducts and acini of the prostate disappear also.

Table 2.1 summarises the Gleason grade of cancer. A lower Gleason grade score indicates a “well differentiated” cancer, with a lower potential to spread. A higher Gleason grade score means a “poorly differentiated” cancer, one that is more likely to spread. The Gleason grade may be an important factor in the doctors decision to recommend a radical prostatectomy to a patient [11].

### X-ray CT Scan and Bone Scan

As part of the diagnostic process a patient may also be expected to undergo additional tests to determine if cancer has spread beyond the prostate. This may include a X-ray CT\(^3\) scan or a bone scan.

An X-ray CT scan generates a series of images showing cross-sections of the prostate. A bone scan is an imaging technique used to identify the presence of cancer in the bones. A tiny amount of radioactive substance is injected into a patient’s bloodstream and absorbed by the bones. If prostate cancer has spread to the bones, it most often shows up on the scan as a series of “hot spots” along the spine and ribs.

---

\(^3\)CT stands for computerised tomography, also called computerised axial tomography (CAT).
2.1. THE PROSTATE

Figure 2.5: Brachytherapy with ultrasound guidance (Courtesy of The Prostate Centre, University of Toronto, Canada [11]).

2.1.4 Treatment of Prostate Cancer

Patients of prostate cancer have a number of treatment options available, including:

- *Observation* that includes periodic PSA tests and DREs to carefully observe the behaviour of the cancer;

- *Radiation therapy* which is one of the main treatment options available to men with localised prostate cancer and includes external beam and brachytherapy;

- *Surgery* which involves removal of the prostate and surrounding tissues, a portion of the urethra, and the seminal vesicles;

- *Hormone therapy* which is one of the oldest means of treating prostate cancer by depriving the prostate of hormones to reduce the cancer; and

- *Other treatment* such as chemotherapy, cryosurgery (or cryotherapy), high intensity focused ultrasound (HIFU), and transperineal microwave ablation of the prostate (TAP), currently being developed at the University of Toronto.

Each option has its own associated risks and benefits. Refer to [11] for more details. We detail brachytherapy below since it involves ultrasound.

Brachytherapy is a radiotherapy approach which involves placing radioactive seeds directly in the prostate where they remain permanently. This treatment is usually offered to those patients who have tumours with a Gleason score of 6 or less, a PSA of less than 10 and a prostate volume of less than 50 cubic centimetres. Usually 90 to 120 seeds are inserted uniformly throughout the prostate gland. The seeds will deliver a high dose of radiation locally to the prostate.

Before implant of radioactive seeds patients must undergo a mapping of the prostate to allow the brachytherapy team to work out all the technical details of the implant.
A transrectal ultrasound (TRUS) is employed to accomplish this task, and also during implant of the radioactive seeds (see Figure 2.5). The ultrasound images are entered into a planning computer and a plan is generated by the physicist and approved by the radiation oncologist.

It takes approximately 6 months for the brachytherapy seeds to deliver the full dose of radiation. For several months following this treatment, a patient will be monitored with PSA tests on a regular basis. A steadily declining PSA is a good indicator that the treatment has been successful in killing the cancerous cells.

2.1.5 Summary

In this subsection we have reviewed the medical background on the prostate and prostatic disease, especially prostate cancer. As we have seen, transrectal ultrasound (TRUS) plays an important role both in diagnosis and in treatment of prostate cancer. In next section we introduce B-mode ultrasound imaging with emphasis on its signal-processing aspects.

2.2 B-mode Ultrasound Imaging

In the form of a gray-scale (or pseudo-colour) image B-mode ultrasound produces a two-dimensional representation of ultrasound echoes, which reveals a tomographic section of the tissue being scanned. The “strength” of the echoes is denoted by brightness\(^4\) of image pixels on a display, as opposed to A-mode scan lines, which simply present waveforms or amplitudes\(^5\) of the echoes. A block diagram of a B-mode ultrasound imaging system is shown in Figure 2.6.

An image is produced by a B-mode ultrasound instrument as follows (refer to Figure 2.6). First the pulse generator emits a voltage burst or impulse, which passes through the transmitter/receiver to excite the piezoelectric element(s) of the transducer. Electric energy is converted into ultrasound energy by the transducer, which transmits ultrasonic pulses. The ultrasound echoes of the tissue being scanned are then picked up by the transducer, which converts the received echoes back into electric voltage bursts.

The received radio frequency echo signal then passes through the transmitter/receiver to reach the radio frequency amplifier, coupled with time gain compensation (TGC). The amplified and compensated signal then goes through the demodulator, which picks up the envelope of the signal from the carrier waves. The envelope of the signal is then subject to further amplification before it reaches the digital scan converter (DSC), which is essentially a memory matrix capable of storing the echoes together with the information of their positions (time of arrival and direction). The data stored in the DSC are then read out in a television raster format and fed as a video signal to the (gray-scale or

\(^4\)Hence the “B” in the term B-mode or B-scan.

\(^5\)Hence the “A” in the term A-mode.
2.2. B-MODE ULTRASOUND IMAGING

![Diagram of B-mode ultrasound imaging system]

Figure 2.6: A block diagram of a B-mode ultrasound imaging system (adapted from [16, 18]).

colour) display. The time logic module coordinates timing of the pulse generator, time gain compensation, and the digital scan converter.

In the following text, we focus on the signal-processing aspects of a ultrasound imaging system. Namely, we are concerned about the details of the signal processing procedures that are applied to the returns from the transducer to produce a B-mode ultrasound image. For the physics aspects of ultrasonics refer to [12, 13]. Refer to [14–18] for introductions on medical (clinical or diagnostic) applications of ultrasound. Refer to [19] for a review on the current status and future prospect of ultrasonic imaging.

2.2.1 Pre-processing

The signals reaching a transducer as a result of echoes from tissue (or tissue boundaries) are very weak, usually in the microvolt and millivolt range. So it is necessary to amplify them to a voltage range suitable for further processing, such as sampling and analog-to-digital (AD) conversion. The amplifier can either be linear or nonlinear. With a linear amplifier, when the overall gain is too high, the large echoes will be presented very strongly, which may result in saturation of the display system, and many small echoes be overemphasized, which also implies that the possibility of the noise interference will increase. On the other hand, too low a gain may result in very few echoes being displayed. For this reason, amplifiers are often of the logarithmic form and, therefore favor weak echoes from small internal interfaces such as liver, kidney and prostate, more than strong ones from large boundary interfaces such as soft tissue/bone, fat/muscle.

Another reason of pre-processing is to reduce the dynamic range of the echoes received by the transducer, which is of the order of 100 dB and in excess of the dynamic ranges of both the display system and human eyes, which are of the order of around 30 dB.
Thus logarithmic amplifiers also serve a signal compression purpose, i.e., they reduce the dynamic range of the echo signals, because the large signals are compressed into a much smaller range due to the logarithmic characteristics of the amplifiers.

More elaborate processing may also be incorporated into the process of pre-amplification to emphasize the leading edges of the echo signal. For instance, a differentiator can be employed to produce pulses whose heights correspond to the slope of the signal edges. The pulses are then combined with the amplified signal to give an echo signal with enhanced edges.

2.2.2 Noise Suppression by Thresholding

Within the echoes of extremely low amplitude there is always a significant quantity of electronic noise and unwanted echo information which can be rejected by introducing a threshold for echo amplitude. This operation also reduces the dynamic range by at least 5dB. The level of the threshold is usually set by the manufacturer.

2.2.3 Time Gain Compensation (TGC)

Time gain compensation (TGC) is also referred to as swept control or depth gain compensation (DGC) in the literature. The echoes returning from deep within the tissue are much weaker than those from the more superficial structures. This is attributed to attenuation of the ultrasound pulse and echoes within the patient. Not only is the transmitted pulse reduced in strength as it travels deeper and deeper within the patient, but the reflected echoes generated by this pulse themselves also undergo the same attenuation process as they return to the transducer.

The attenuation of an ultrasound pulse is the sum of all forms of energy loss as it passes through tissue, most of which result from reflection, scattering, refraction, absorption and wavefront divergence. If no action were taken to compensate for this, the echoes from distant structures would be too weak to be detectable within the ultrasound image. It is, therefore, necessary to compensate for this attenuation and the mechanism by which this is achieved is called time gain compensation (TGC) or depth gain compensation (DGC).

This attenuation is approximately exponential, therefore the TGC is usually made to vary with time as an exponential function and if displayed in logarithmic scale, is thus a straight line of certain slope. The characteristics of TGC may be further controlled by changing several parameters, such as the starting level of the TGC, the rate of increase of gain, i.e., the slope if in logarithmic scale, and TGC delay, which determines the time delay of the starting point of the slope after the instant of transmission. Adaptive gain compensation (AGC) has been introduced by some manufacturers, which senses the general decrease in echo amplitude with depth and applies the appropriate compensation. It is believed that its performance is close to that of an experienced operator.
2.2.4 Demodulation (Rectification and Envelope Detection)

The pre-amplified and time gain compensated signal is still a high frequency, full wave radio frequency (RF) signal. The positive and negative oscillations of the wave do not necessarily represent true structure of the tissue along the ultrasound beam. However, the major fluctuations in the outline of the signal do bear some relationship to the anatomical structures. So the next step is to rectify the signal, namely, remove all the negative components of the signal. The rectified signal is then further envelope detected, eliminating the RF component and leaving only the outline. This is equivalent to low pass filtering. Refer to section 4.1.3 (page 56) for more details.

2.2.5 Digital Scan Conversion

Digital scan conversion, which allows gray-scale images to be displayed on a CRT monitor, typically accepts information in one scan mode, for example, scan line by scan line from a sector scanner, and stores it electronically, processes it and sends it to the display unit in a second scan pattern, for example, as a series of horizontal television-raster lines. The format of incoming information is determined by the design of the scanner. For sector scanners the information is written in as a series of oblique lines and, as the image line is at an angle to columns and rows, it passes obliquely through more than one element in some rows. In this situation an identical value is stored in the adjacent elements and the original straightness of the data line is degraded. In practice, the size of each picture element is very small and these slight excursions are barely detectable in the final image.

2.2.6 Post-processing

Post-processing may be done at the output of the digital scan converter, such as light-dark reversal, and gray scale enhancement. Unlike pre-processing, several post-processing options can be tried (without rescanning the patient) by reversion to the unprocessed data since it involves manipulation of data after it is stored in memory. Image improvement procedures such as frame averaging or interpolation can also be regarded as post-processing steps.

2.2.7 Summary

In summary, described above are the general signal processing steps that are usually employed to process the echo signal received from the transducer prior to gray-scale presentation, including noise suppression and pre-processing, time gain compensation, demodulation, digital scan conversion, and post-processing. However, full details of these signal processing procedures are rarely revealed by the manufacturers. In section 4.1 we present a hypothetical ultrasonic imaging system for the purpose of simulation of B-mode ultrasound images in Chapter 4 and ultimately for speckle reduction in Chapter 6.
2.3 The Proscan Plus Imaging System

The ultrasound images of the prostate used in this study were collected from a Proscan Plus multi-frequency/multi-view prostate imaging system, which is manufactured by Carolina Medical Inc. and “specially designed to provide clear, high resolution images of the entire prostate as well as the scrotum” [20]. Figure 2.7(a) illustrates the Proscan Plus imaging system and its accessories, with the main console including the CRT display on the far right. In front of it are a keyboard and a trackball. Shown in the foreground are two transrectal probes, behind which is a transrectal biopsy needle guide.

This system features a 5.0/7.5 MHz single crystal multi-view transrectal probe, which can produce consistent high quality images of the prostate gland on the transverse, longitudinal, and end-fire planes. Figure 2.7(b) illustrates these three imaging planes. The transverse images have a view angle of 166°, and manifest prostate gland symmetry, volume, and capsular margin. Longitudinal images are used for the examination of the bladder neck, seminal vesicles and prostatic apex, and have a view angle of 150°. A 120° end-fire view of the prostate gland can be produced by slightly rotating the probe.

The multi-frequency capability of the system lies in the variability of the frequency of the ultrasound transducer (probe), which enables the operator to examine the posterior
2.4. APPEARANCE OF ULTRASOUND IMAGES OF THE PROSTATE

<table>
<thead>
<tr>
<th>MAIN SYSTEM</th>
<th>TRANSRECTAL PROBE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Display</td>
<td>9 in/22.8 cm CRT</td>
</tr>
<tr>
<td>Display Standard</td>
<td>EIA, 525 lines</td>
</tr>
<tr>
<td>Gray scale</td>
<td>128 levels</td>
</tr>
<tr>
<td>Dynamic Range</td>
<td>48 dB</td>
</tr>
<tr>
<td>Gain Range</td>
<td>0-75 dB</td>
</tr>
<tr>
<td>Display Matrix size</td>
<td>640×480 pixels</td>
</tr>
<tr>
<td>Memory Matrix size</td>
<td>1024×512 pixels</td>
</tr>
<tr>
<td>Transverse View Angle</td>
<td>166°</td>
</tr>
<tr>
<td>Longitudinal View Angle</td>
<td>150°</td>
</tr>
<tr>
<td>Frontal View Angle</td>
<td>120°</td>
</tr>
<tr>
<td>Frequency</td>
<td>5.0/7.5 MHz</td>
</tr>
<tr>
<td>Focal Zone</td>
<td>16-41 mm</td>
</tr>
<tr>
<td>Axial Resolution</td>
<td>1mm</td>
</tr>
<tr>
<td>Lateral Resolution</td>
<td>1mm</td>
</tr>
</tbody>
</table>

Table 2.2: Specifications of the Proscan Plus imaging system (Courtesy of Carolina Medical Inc. [20]).

of the prostate in the high frequency (7.5 MHz) and obtain a clearer view of the anterior by switching to the low frequency (5.0 MHz). The split screen function is another feature of this system, which allows two views (for instance, transverse and longitudinal) to be displayed simultaneously on the display. Table 2.2 summarises the specifications of the Proscan Plus imaging system including the transrectal transducer.

One of the accessories of the system is a VHS video cassette recorder, via which the examination process can be recorded on a video cassette tape. Such video clips were digitised on an SGI O2 workstation resulting in sequences of ultrasound images of the prostate used in this study.

2.4 Appearance of Ultrasound Images of the Prostate

In this section we attempt to explain the appearance of B-mode ultrasound images of the prostate. We try to summarise the correspondence, if any, between the grayscale presentation and the underlying anatomical or pathological structures in the prostate. Concretely, we attempt to answer the nontrivial question: what tissues are, or what makes tissues hypoechoic or hyperechoic? For more details refer to [2,3], which provide a thorough and professional treatment on this issue.

The normal peripheral and central zones of the prostate have a relatively homogeneous texture, corresponding to the uniform distribution of glands in these area. The transition zone, the locus of benign prostatic hyperplasia (BPH), has a more heterogeneous pattern, reflecting the variation in the size and distribution of glands. Muscular, stromal, or fibrous tissue free of glands, such as the periurethral smooth muscle or the ejaculatory duct complex, has few interfaces and appears hypoechoic. The lumen of a hollow structure has no interfaces to reflect the ultrasound waves and appears anechoic. Due to the low level acoustic reflectivity from the smooth muscle surrounding the ejaculatory ducts, it may appear hypoechoic.

The hypoechoic appearance on ultrasound may also result from the pathologic features of prostate cancer. Most cancers consist of a dense mass of cells differing in structure from the normal prostate gland. Cancer destroys the normal glandular structure, replacing it...
with a packed mass of cells either in form of small glands or of a sheet without discernible structure. The malignant tissue contains few sonographically detectable interfaces and therefore appears hypoechoic in relation to the adjacent normal tissue. The most common feature of virtually all visible cancer is a central region hypoechoic relative to the peripheral zone of the normal prostate. High-grade tumors do not form glandular structures and completely destroy the normal prostate structure by replacing it with a mass of cells, and consequently are more likely to appear hypoechoic or even anechoic. Figure 2.8 compares the ultrasound images of a normal prostate gland and that with cancer.

The outer gland, in the normal young man, who has little or no enlargement of the inner gland (due to the absence of BPH), has little difference in echogenicity from the normal inner gland. However, in the older man, who frequently has inner gland enlargement (due to BPH), the outer gland becomes compressed. The compression of normal tissue of the outer gland causes a homogeneous increase in the echogenicity, thus a hypoechoic lesion in the compressed, relatively hyperechoic outer gland tissue is more obvious. These acoustic features ...
2.4. APPEARANCE OF ULTRASOUND IMAGES OF THE PROSTATE

![Image](a)

![Image](b)

**Figure 2.9:** A sample of the ultrasound images of the prostate used in this study. (a) The original image “us1” showing a transverse section of the prostate. (b) After non-image portions are cropped.

reflectivity changes, or acoustic mismatching in adjacent tissues, allow differentiation of one tissue from another.

“Acoustic shadowing” is the term used to describe the pattern in which the majority of the sound waves are returned by reflection from a dense echogenic barrier. Few ultrasound waves are transmitted beyond the barrier, and imaging distal to it is inadequate. In prostate ultrasound, the most common cause of acoustic shadowing is the ring of echo-
dense corpora amylacea or calculi at the margin of the transition zone. Another common cause of shadowing is gas within the rectum.

Some lesions are more echogenic than normal or residual prostate tissue and appear hyperechoic. They can vary from subtle areas of increased echogenicity to areas that are brightly and thickly echogenic, which usually represent prostatic calculi or corpora amylacea. Rarely are tumors completely hyperechoic, although some cancers contain hyperechoic foci.

2.5 Sample Ultrasound Images of the Prostate

Figure 2.9(a) shows a sample of the original digitised images, which presents a transverse section of the prostate enclosed by a rectangular axis with calibration markers. Along with this image are displayed other textual information, such as that on patient and physician on top (removed for protection of privacy) and help menus for keyboard shortcuts at the bottom. These non-image portions are irrelevant to the study in this thesis and were then removed only leaving the image of the prostate, which is shown in Figure 2.9(b). Image “us1” will be used for demonstration of speckle distribution modelling in Chapter 4, our speckle reduction scheme in Chapter 6, and contour extraction scheme in Chapter 10.

Figure 2.10 illustrates a longitudinal image of the prostate, which is known as “us5” in the following text and used for demonstration of the speckle reduction scheme in Chapter 6. The split screen function of the Proscan Plus imaging system is demonstrated in Figure 2.11 in which the left panel shows a longitudinal image and the right panel a transverse image. These two images are respectively referred to as “us6a” and “us6b”, and used for demonstration of speckle distribution modelling in Chapter 4. Figure 2.12 shows two more sample ultrasound images of the prostate. Image “vs01” (Figure 2.12(a)) will be used in Chapter 10 for demonstration of our contour extraction scheme and image “vs04” (Figure 2.12(b)) will be used in Chapter 7.

This concludes our brief overview of the scientific and medical background on the prostate and ultrasound of the prostate. We continue the survey in Chapter 3 on speckle distribution models, and in Chapter 5, where we review methods for suppressing speckle. We analyse in Chapter 8 the current literature on the problem of delineating the contour of the prostate in B-mode ultrasound images, after a pilot study in Chapter 7, where we explore this problem using the Laplacian of Gaussian (LoG) operator.

---

*In the following text all the ultrasound images of the prostate involved are cropped version of the originals unless explicitly stated otherwise.*
2.5. SAMPLE ULTRASOUND IMAGES OF THE PROSTATE

Figure 2.10: Sample ultrasound image “us5” showing a longitudinal section of the prostate.

Figure 2.11: Split screen showing transverse and longitudinal prostate images. The left panel shows a longitudinal image and the right panel a transverse image.
Figure 2.12: Two more sample ultrasound images of the prostate. (a) Image “vs01”. (b) Image “vs04”.

(a)

(b)
2.5. SAMPLE ULTRASOUND IMAGES OF THE PROSTATE
Chapter 3

Speckle Distribution Models: An Overview

Speckle refers to the granular appearance that compromises the image quality and resolution in optics, synthetic aperture radar (SAR), and ultrasound. Due to the existence of speckle the appearance of a B-mode ultrasound image does not necessarily relate to the internal structure of the object being scanned.

This distracting artifact is a result of the coherent nature of the imaging process/system. In particular, ultrasound speckle results from the coherent interference of the random backscattered echo signal from within the resolution cell of the ultrasound imaging system. Intuitively, if interference is mainly constructive, the corresponding pixel in the B-scan image will be bright. On the contrary, destructive interference results in a pixel of low gray scale value. Speckle has been treated extensively in the literature for laser, SAR and ultrasound images [21, 22]. Some speckle distribution models will be introduced in section 3.1.

Methods for parameter estimation of these distribution models are introduced in section 3.2, including the method of maximum likelihood estimation (MLE), fitting the probability density function (PDF), fitting the cumulative distribution function (CDF) and the method of moments. These methods allow us to analyse the statistical properties of simulated B-scan images, as described in the next chapter.

Section 3.3 presents the goodness-of-fit test procedures, including the chi-squared test and the Kolmogorov-Smirnov test, which will be used in the next chapter to test our hypotheses and verify whether the pixel values of simulated images really follow various specified distribution models for different scenarios.

3.1 Overview of Speckle Distribution Models

Many statistical distribution models have been proposed in an attempt to characterise the statistical properties of speckled images, including the Rayleigh [21–27], Rician [23, 25, 27–30], K [30–37], Fisher-Tippett [27, 38–40], Nakagami [28, 41–48], Weibull [49–51], and
the generalised Gamma distribution [51, 52]. A thorough treatment of most of the above-
mentioned distribution models can be found in [27, 51]. Apart from these distributions, we
also empirically propose the generalised Fisher-Tippett distribution, which will be used to
analyse the ultrasound images of the prostate in section 4.3.

3.1.1 The Rayleigh Distribution

In a pulse-echo ultrasound imaging system, it is assumed that there are numerous scat-
ters within one resolution cell of the system under the discrete scatterer model (see
section 4.1). These randomly located scatterers reflect spherical wavelets with random
amplitudes and phases. At a particular instant of time, the echo signal received at the
transducer is a sum of the wavelets backscattered from the scatterers in the resolution cell
at a depth corresponding to that instant of time

\[ A = \sum_{i=1}^{N} a_i = X e^{j\phi} \]  

(3.1)

where \( j = \sqrt{-1} \) and \( a_i \) is the complex echo signal from the \( i \)-th scatterer [26, 27]. \( N \) is
the number of scatterers in the resolution cell. \( X \) is the net magnitude and \( \phi \) is the net
phase.

Under the assumptions that

1. the number of scatterers in one resolution cell of the imaging system is large and
close to one another (that is, their size is far less than the ultrasound pulse width,
thus “unresolvable” to the emitted ultrasound pulse);

2. the locations of the scatterers are random and independent of one another;

3. the phases of the backscattered wavelets from each scatterer are random, indepen-
dent of one another, and uniformly distributed over (0, 2\( \pi \))

the Central Limit Theorem applies and accumulation of the wavelets is two-dimensional
random walk problem, with the wavelets backscattered from each scatterer representing a
step of random length and direction in the complex plane [21–26]. The resultant complex
amplitude \( A \) is a complex Gaussian random variable with independent and identically
distributed real and imaginary parts and its probability density function follows a circular
Gaussian function. The magnitude \( X \) follows a Rayleigh distribution [23]

\[ p_X(x) = \frac{x}{\sigma^2} e^{-x^2/2\sigma^2} \]  

(3.2)

where \( x \geq 0 \) and \( \sigma > 0 \).

In keeping with common convention [21, 24], Equation (3.2) can also be written as

\[ p_X(x) = \frac{x}{\psi} e^{-x^2/2\psi} \]  

(3.3)
Figure 3.1: The Rayleigh distribution for different parameter values.

where $\psi = \sigma^2$. Figure 3.1 illustrates the probability density function of the Rayleigh distribution for different parameter values. As we can see, the smaller the parameter $\sigma$, the more right skewed the Rayleigh distribution.

The $\nu$-th moment of $X$ about the origin (the raw moment), $\mu_{\nu}$, is given by

$$\mu_{\nu} = E[X^{\nu}] = \int_0^\infty x^{\nu} p_X(x) \, dx. \quad (3.4)$$

where $E[\cdot]$ stands for expectation. For the Rayleigh distribution, the first two moments are respectively

$$\mu = \sqrt{\frac{\pi}{2}} \sigma = \sqrt{\frac{\pi\psi}{2}} \quad (3.5)$$

and

$$\mu_2 = 2\sigma^2 = 2\psi. \quad (3.6)$$

According to [21, 24] the signal-to-noise ratio (SNR) of a speckled image is defined as

$$\text{SNR} = \frac{\mu}{\sqrt{\mu_2 - \mu^2}} \quad (3.7)$$

and the inverse of SNR is called the speckle contrast. Therefore, for the Rayleigh distribution

$$\text{SNR} = \sqrt{\frac{\pi}{4 - \pi}} \approx 1.9131. \quad (3.8)$$
3.1. OVERVIEW OF SPECKLE DISTRIBUTION MODELS

In practice the SNR of a B-scan image is computed as the ratio of the mean and standard deviation of the pixel values of the image. If the SNR of an image is approximately 1.9131, the speckle is called “fully developed” (or “fully formed”). Thus the SNR of an image can be used to detect fully developed speckle [21].

The assumptions for the Rayleigh distribution rarely hold true at the same time in reality, therefore many non-Rayleigh distribution models have been proposed to deal with these situations.

3.1.2 The Rician Distribution

When there are periodically located scatterers or other nonrandom clustered components in the medium, the amplitude of the pixels of the resultant B-scan image will deviate from the Rayleigh statistic to become Rician \(^1\) distributed [25,27–29]. The Rician distribution is defined by

$$p_X(x) = \frac{x}{\sigma^2}I_0\left(\frac{s^2x}{\sigma^2}\right)e^{-\left(x^2+s^2\right)/2\sigma^2}$$

(3.9)

where \(x \geq 0\), \(\sigma > 0\) and \(s \geq 0\). \(I_0(\cdot)\) is the modified Bessel function of the first kind of order 0 [50,53]. \(s^2\) characterises the backscattered energy from the nonrandom components and \(\sigma^2\) represents the diffuse energy from the random scatterers. The SNR for Rician distributed speckle is no less than 1.9131 [27].

\(^1\)Named after R.O. Rice, who proposed this distribution in his analysis of random noise [23]. It is also referred to as the Rice or Ricean distribution in the literature.
When $s = 0$ it becomes the Rayleigh distribution. Figure 3.2 shows the probability density function of the Rician distribution of various parameters. As we can see, generally the smaller the parameter $s$ (or $\sigma$), the more right skewed the Rician distribution. Note the special case of a Rayleigh distribution when $s = 0$ and $\sigma = 0.25$ (light colour solid line). The Rician distributions are all post-Rayleigh, that is, their modes are always to the right of that of the Rayleigh distribution.

The even order moments of the Rician distribution have a closed-form expression [27]. Especially the second and fourth moments respectively are

$$\mu_2 = 2\sigma^2 + s^2$$  \hspace{1cm} (3.10)

and

$$\mu_4 = 8\sigma^4 + 8\sigma^2 s^2 + s^4.$$ \hspace{1cm} (3.11)

Thus the parameter $\sigma$ and $s$ can be estimated from the second and fourth moments of the sample data using the method of moments, to be described in the next section.

### 3.1.3 The K Distribution

The K distribution was first proposed by Jakeman and Pusey in their study on sea echo of microwave radar [31]. It has proven to be a promising model in the context of SAR and ultrasonic imaging [27, 32–37]. It was shown that when the scatterer density is sparse, the K distribution provides a reasonable model for the statistics of speckle [27].

Speckle is said to be K-distributed if its probability density function is given by

$$p_X(x) = \frac{2b}{\Gamma(\alpha)} \left( \frac{bx}{2} \right) ^{\alpha} K_{\alpha-1}(bx)$$  \hspace{1cm} (3.12)

where $x \geq 0$, $b > 0$ and $\alpha > 0$. $\Gamma(\cdot)$ is the well known Gamma function and $K_{\alpha-1}(\cdot)$ is a modified Bessel function of the second kind of order $\alpha - 1$ [50,53]. Parameter $\alpha$ is related to the number of scatterers per resolution element. Parameter $b$ is given by

$$b = 2\sqrt{\frac{\alpha}{E[X^2]}}$$  \hspace{1cm} (3.13)

The K distribution approaches the Rayleigh distribution as $\alpha$ tends to $\infty$. The SNR for K distributed speckle is less than 1.9131 [27]. Figure 3.3 shows the probability density function of the K distribution of various parameters, together with a Rayleigh distribution (light colour solid line) with the parameter $\sigma = 0.25$ as a “reference”. As we can see, compared to the Rayleigh distribution, the K distribution is more right-skewed, thus becomes pre-Rayleigh.

The moments of the K distribution have a closed-form expression [27, 31, 33–35, 40].
3.1. OVERVIEW OF SPECKLE DISTRIBUTION MODELS

Figure 3.3: The K distribution of various parameters.

Especially, the normalised fourth moment is given by

\[
\frac{\mu_4}{\mu_2^2} = 2 \left( 1 + \frac{1}{\alpha} \right)
\]

and the normalised sixth moment is given by

\[
\frac{\mu_6}{\mu_2^3} = 6 \left( 1 + \frac{1}{\alpha} \right) \left( 1 + \frac{2}{\alpha} \right).
\]

Thus the parameter \( \alpha \) can be estimated from the sample data using the method of moments, to be described in the next section, and \( b \) can then be estimated using Equation (3.13).

3.1.4 The Fisher-Tippett Distribution

The Fisher-Tippett distribution is also referred to as the extreme value distribution or log-Weibull distribution [50]. Kaplan and Ma proposed a distribution of the Fisher-Tippett form\(^2\) to model the statistical properties of logarithmic compressed B-mode images in [38, 50].

\(^2\)Note this distribution is slightly different from the original Fisher-Tippett distribution defined in [50] in that it is actually left-and-right “flipped” (about the vertical axis). Nevertheless, it is still referred to as the Fisher-Tippett distribution in the following text.
Figure 3.4: The Fisher-Tippett distribution of various parameters.

The Fisher-Tippett distribution is given by

\[ p_X(x) = \frac{1}{\beta} e^{-(\alpha-x)/\beta - e^{-(\alpha-x)/\beta}} \]  

which can also be written as

\[ p_X(x) = \frac{1}{\beta} \exp(-g - \exp(-g)) \]  

where

\[ g = \frac{\alpha - x}{\beta}. \]  

The first moment about the origin is given by

\[ \mu = \alpha - \gamma \beta \]  

where \( \gamma \approx 0.5772 \) is the Euler-Mascheroni constant [50]. The second moment about the mean (namely the variance) is given by

\[ \mu'_2 = E[(X - E[X])^2] = \frac{1}{6}(\pi \beta)^2 \]  

On the basis of this model, Dutt\(^3\) and Greenleaf proposed an adaptive speckle reduction

\(^3\)Note there was an error in equation (7.5) in [27] and equation (8) in [40]. The author inadvertently missed a minus sign between \(-g\) and \(\exp(-g)\).
3.1. OVERVIEW OF SPECKLE DISTRIBUTION MODELS

filter for logarithmic compressed B-mode ultrasound images [27, 40]. This model was also used for robust estimation of ultrasound pulses, which usually arises as a part of the problem of ultrasound image restoration [55].

Figure 3.4 shows the probability density function of the Fisher-Tippett distribution of various parameters. It is quite apparent $\alpha$ is a displacement parameter and $\beta$ is a scaling parameter which controls the spread of the distribution (refer to Equation (3.18)). Note that all these density functions are negative (left) skewed, which does not agree with the histograms of our ultrasound images of the prostate (see Figure 4.19-4.22). To be consistent with the real ultrasound images, an ad hoc logarithmic compression is introduced, to be described in section 4.1.4.

Nevertheless, study of the Fisher-Tippett distribution has led us to propose the generalised Fisher-Tippett distribution, as described later, which will be used in the analysis of statistical properties of the ultrasound images of the prostate in section 4.3.

3.1.5 The Nakagami Distribution

Nakagami first proposed the distribution that bears his name in the context of high-frequency radio propagation [41]

$$p_\chi(x) = \frac{2m^m x^{2m-1}}{\Gamma(m)\Omega^m} e^{-mx^2/\Omega}$$

(3.21)

where $\Gamma(\cdot)$ is the Gamma function. $m \geq 0.5$ is the shape parameter and $\Omega > 0$ is a scaling factor.

This distribution is analytically simple and straightforward compared with the Rician and K distribution. It was recently applied to ultrasonic B-mode imaging as a general statistical model for tissue characterisation and classification [42–46]. It was demonstrated that the Nakagami distribution can account for a wide range of scattering conditions that exist in tissue, and was able to model the backscattered signal from tissues such as the human breast. Shankar showed that parameter $m$ could be relaxed to be less than the conventional value 0.5 in the context of ultrasonic imaging [42].

The $\nu$-th moment about the origin of the Nakagami distribution [41, 47, 48] is given by

$$\mu_\nu = E[X^\nu] = \frac{\Gamma(m + \frac{\nu}{2})}{\Gamma(m)} \left( \frac{\Omega}{m} \right)^{\frac{\nu}{2}}$$

(3.22)

and especially the second moment is

$$\mu_2 = \frac{\Gamma(m + 1)}{\Gamma(m)} \left( \frac{\Omega}{m} \right) = \Omega$$

(3.23)

4Noting that the Gamma function has the property $\Gamma(x) = (x-1)\Gamma(x-1)$ [50].
and the fourth moment is

$$
\mu_4 = \frac{\Gamma(m + 2)}{\Gamma(m)} \left( \frac{\Omega}{m} \right)^2.
$$

(3.24)

It is easily derived from the above two equations that

$$
m = \frac{\mu_2^2}{\mu_4 - \mu_2^2}.
$$

(3.25)

When $m = 0.5$ the Nakagami distribution becomes a one-sided Gaussian distribution [48]. The Rayleigh distribution is a special case of the Nakagami distribution when $m = 1$. When $m > 1$, the Nakagami distribution approaches the Rician distribution [51]. It can be shown that as $m$ increases, the probability density function of the Nakagami distribution becomes asymptotically equivalent to that of the log-normal distribution\(^5\) [28,41].

Therefore, the Nakagami distribution is able to encompass a variety of distributions. Figure 3.5 illustrates the probability density function of the Nakagami distribution of various parameters. Note the special case of the one-sided Gaussian ($m = 0.5$) and Rayleigh distribution ($m = 1.0$).

\(^5\)The log-normal distribution [50,56] is defined by

$$
p_X(x) = \frac{1}{\sqrt{2\pi\sigma_x}} e^{-(\ln x - \mu)^2/2\sigma^2}
$$

(3.26)

where $\mu = E[\ln(X)]$ and $\sigma^2 = \text{var}[\ln(X)]$ is the variance of $\ln(X)$. 
3.1. OVERVIEW OF SPECKLE DISTRIBUTION MODELS

3.1.6 The Weibull Distribution

The Weibull distribution is well known in the context of reliability and lifetime modeling [49, 50]. Its probability density function is defined by

\[ p_X(x) = \frac{\beta}{\alpha^\beta} x^{\beta-1} e^{-(x/\alpha)\beta} \]  

(3.27)

where \( \alpha > 0 \) controls the spread of the distribution and \( \beta > 0 \) is a shape parameter. It was recently used in the analysis of the statistics of envelope of high-frequency ultrasonic backscatter signal from human skin in vivo [51].

The first and second moment about the origin are respectively

\[ \mu = E[X] = \alpha \Gamma \left( 1 + \frac{1}{\beta} \right) \]  

(3.28)

and

\[ \mu_2 = E[X^2] = \alpha^2 \Gamma \left( 1 + \frac{2}{\beta} \right) \]  

(3.29)

From the above two equations the following function can be constructed

\[ f(\beta) = \frac{\mu_2}{\mu^2} = \frac{\Gamma \left( 1 + \frac{2}{\beta} \right)}{\Gamma^2 \left( 1 + \frac{1}{\beta} \right)} \]  

(3.30)

and it can be shown that \( f(\beta) \) is a monotonic function of \( \beta \) [51]. Figure 3.6 shows the plot of \( f(\beta) \) for 0.5 \( \leq \beta \leq 3.0 \), which confirms that \( f(\beta) \) is a monotonic decreasing function.

The SNR is given by

\[ \text{SNR} = \frac{\mu}{\sqrt{\mu_2 - \mu^2}} = \frac{\Gamma \left( 1 + \frac{1}{\beta} \right)}{\sqrt{\Gamma \left( 1 + \frac{2}{\beta} \right) - \Gamma^2 \left( 1 + \frac{1}{\beta} \right)}} \]  

(3.31)

and it can be shown that the SNR is a monotonic increasing function of \( \beta \) and can take any positive value [51]. Thus the Weibull distribution is capable of modelling a variety of conditions, both pre-Rayleigh and post-Rayleigh. Figure 3.7 illustrates the probability density function of the Weibull distribution of various parameters.

3.1.7 The Generalised Gamma Distribution

The generalised Gamma distribution was first proposed in [52] and its probability density function is given by

\[ p_X(x) = \frac{b}{a \Gamma(\nu)} \left( \frac{x}{a} \right)^{b-1} e^{-\left( \frac{x}{a} \right)^b} \]  

(3.32)

where \( a > 0 \) is a scale parameter and \( b, \nu > 0 \) are shape parameters. The generalised Gamma distribution encompasses several distributions as special cases including Gamma (when \( b = 1 \)), Rayleigh (when \( b = 2 \) and \( \nu = 1 \)), exponential (when \( b = 1 \) and \( \nu = 1 \)), and Weibull (when \( \nu = 1 \)) distribution [51].
Figure 3.6: A plot of the function $f(\beta)$ of relevance to the Weibull distribution.

Figure 3.7: The Weibull distribution of various parameters.
Figure 3.8: The generalised Gamma distribution of various parameters.

Figure 3.8 illustrates the probability density function of the generalised Gamma distribution of various parameters. As we can see, like the Weibull and Nakagami distribution, the generalised Gamma distribution is capable of modelling a variety of conditions, both pre-Rayleigh and post-Rayleigh due to the flexibility provided by the two shape parameters.

This distribution was recently applied in the context of high-frequency ultrasound and found to fit the backscattered envelope signal from skin tissues better than the other distributions in comparison [51].

3.1.8 The Generalised Fisher-Tippett Distribution

Since the Fisher-Tippett distribution is not compatible with the histograms of our ultrasound images of the prostate (compare Figure 3.4 with Figure 4.19-4.22), we empirically propose the generalised Fisher-Tippett distribution.

We first “flip” it about the vertical line $x = 0.5$ to restore its right skewness, which is a matter of change of the independent variable $x = -x' + 1$. Thus Equation (3.18) becomes

$$g(x') = \frac{\alpha + x' - 1}{\beta}$$  \hspace{1cm} (3.33)
which after a change of symbol is equivalent to

\[ g = \frac{\alpha + x - 1}{\beta}. \]  

(3.34)

We further introduce a scaling parameter \( c \) to fine-tune its shape making it more flexible, and Equation (3.17) becomes

\[ p_X(x) = \frac{1}{\beta} \exp(-cg - \exp(-g)). \]  

(3.35)

Initially \( c \) is set to 1. Thus Equation (3.19) becomes

\[ \mu = \gamma\beta - \alpha + 1 \]  

(3.36)

and the variance in Equation (3.20) remains unchanged.

Figure 3.9(a) illustrates the probability density function of the generalised Fisher-Tippett distribution under the effect of parameters \( \alpha \) and \( \beta \) (parameter \( c \) is fixed at \( c = 1.0 \)). Setup of parameter \( \alpha \) and \( \beta \) is the same as that in Figure 3.4. As we can see, the p.d.f. curves in Figure 3.9(a) are a “flipped” (about \( x = 0.5 \)) version of that in Figure 3.4, as expected. Figure 3.9(b) shows the p.d.f. curves under variation of parameter \( c \) (parameter \( \alpha \) and \( \beta \) are fixed at \( \alpha = 0.8 \) and \( \beta = 0.1 \)).

### 3.2 Parameter Estimation

In section 4.2 we need to analyse the statistical properties of simulated B-scan images in different situations. For instance, we would like to check whether the pixel values of a simulated B-scan image of a homogeneous phantom follow the Rayleigh distribution. This involves parameter estimation and goodness-of-fit tests for the above-mentioned probability distribution models. In this section we present methods for parameter estimation including the method of maximum likelihood estimation (MLE), fitting the probability density function (PDF), fitting the cumulative distribution function (CDF), and the method of moments. Goodness-of-fit tests, including the chi-squared test and the Kolmogorov-Smirnov test, will be introduced in the next section.

The **method of maximum likelihood** and the **method of moments** are the two most common techniques for obtaining point estimators. Although estimators produced by the method of moments can be obtained with a minimum of computational efforts, they are only regarded as “reasonable” since they are not necessarily sufficient or efficient. The MLE method usually requires significantly more computation than the method of moments, but is still preferred nevertheless, because ML estimators have certain efficiency properties (they are approximately unbiased) [56] and “are always functions of sufficient statistics” [57].

The following text will be focused on the method of MLE in estimating the distribution
3.2. PARAMETER ESTIMATION

Figure 3.9: The generalised Fisher-Tippett distribution of various parameters. (a) Parameter $c$ is fixed at $c = 1.0$. Setup of parameters $\alpha$ and $\beta$ is the same as that in Figure 3.4. Note the p.d.f. curves are a “mirrored” version of that in Figure 3.4. (b) Effect of parameter $c$ ($\alpha = 0.8$ and $\beta = 0.1$).
parameters, although the method of moments will be employed to produce an initial estimation of the parameters, which is fed into the method of MLE, PDF or CDF for a final solution.

3.2.1 Maximum Likelihood Estimation (MLE)

Maximum likelihood estimation (MLE) is one of the most important methods of statistical inference and is recommended by most statisticians for estimation of parameters. A thorough treatment of MLE can be found in [57].

Let $X_1, X_2, ..., X_n$ represent independent random variables taken from a distribution whose probability density function (or probability mass function for the discrete case) is given by $f_X(x; \theta_1, \theta_2, ..., \theta_k)$, where $\theta_1, \theta_2, ..., \theta_k$ are $k$ parameters of the distribution to be estimated. The joint distribution of the random variables is

$$f_{X_1, X_2, ..., X_n}(x_1, x_2, ..., x_n; \theta_1, \theta_2, ..., \theta_k) = f_X(x_1; \theta_1, \theta_2, ..., \theta_k) \cdot f_X(x_2; \theta_1, \theta_2, ..., \theta_k) \cdot \ldots \cdot f_X(x_n; \theta_1, \theta_2, ..., \theta_k) = \prod_{i=1}^{n} f_X(x_i; \theta_1, \theta_2, ..., \theta_k)$$

(3.37)

The observed sample values $x_1, x_2, ..., x_n$ are considered fixed and the above equation is thought of as being a function of the parameters $\theta_1, \theta_2, ..., \theta_k$, that is,

$$L(\theta_1, \theta_2, ..., \theta_k) = L(\theta_1, \theta_2, ..., \theta_k; x_1, x_2, ..., x_n) = \prod_{i=1}^{n} f_X(x_i; \theta_1, \theta_2, ..., \theta_k)$$

(3.38)

which is referred to as the likelihood function.

The likelihood function $L(\theta_1, \theta_2, ..., \theta_k)$ measures the probability that the observed sample data $x_1, x_2, ..., x_n$ follow a distribution specified by the parameters $\theta_1, \theta_2, ..., \theta_k$. Those values $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ that maximise the likelihood function

$$L(\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k) \geq L(\theta_1, \theta_2, ..., \theta_k) \quad \text{for all} \quad \theta_1 \neq \hat{\theta}_1, \theta_2 \neq \hat{\theta}_2, ..., \theta_k \neq \hat{\theta}_k$$

(3.39)

are called maximum likelihood estimate (MLE) of the parameters. Intuitively, the MLE $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ produce the largest probability of obtaining the observed sample values. In other words, the probability function whose parameters are $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ best explains or agrees most closely with the observed data.

Since the natural logarithm function is monotonic, maximising $\ln L(\theta_1, \theta_2, ..., \theta_k)$ is equivalent to maximising the original likelihood function $L(\theta_1, \theta_2, ..., \theta_k)$. It is usually convenient and a common practice to work with $\ln L(\theta_1, \theta_2, ..., \theta_k)$. Taking the natural logarithm of both sides of equation (3.38) gives

$$\ln L(\theta_1, \theta_2, ..., \theta_k) = \sum_{i=1}^{n} \ln f_X(x_i; \theta_1, \theta_2, ..., \theta_k)$$

(3.40)
3.2. PARAMETER ESTIMATION

If \( L \) or \( \ln L \) is differentiable with respect to \( \theta_1, \theta_2, \ldots, \theta_k \), maximising it is a standard calculus problem, which involves taking partial derivatives of \( L \) or \( \ln L \) with respect to \( \theta_1, \theta_2, \ldots, \theta_k \), setting the partial derivatives to 0 and solving for the unknown parameters \( \theta_1, \theta_2, \ldots, \theta_k \). For the simple cases such as the normal distribution a closed form solution can be obtained. However, for many cases it is impossible to obtain a closed form expression and a solution has to be found numerically [27, 34].

3.2.2 Nonlinear Curve Fitting

Apart from the method of MLE, it is also sensible to examine the problem of parameter estimation from a different point of view, that is, nonlinear curve fitting. The parameters will be estimated implicitly by fitting the probability density function (p.d.f.) of a specified distribution to the histogram of the image in a least-squares sense. In other words, a best fit will minimise the sum of squares due to error (the summed square of residuals), which is defined as

\[
SSR = \sum_{x=0}^{1} |p_S(x) - p_X(x; \theta_1, \theta_2, \ldots, \theta_k)|^2
\] (3.41)

where \( x \) represents the different gray scale values each pixel can possibly take (normalised to [0,1] with 0 as black and 1 as white). \( p_S(x) \) represents the histogram, that is, the observed distribution function of the pixel values of a simulated image. \( p_X(x; \theta_1, \theta_2, \ldots, \theta_k) \) is the p.d.f. of a distribution model specified by the \( k \) parameters \( \theta_1, \theta_2, \ldots, \theta_k \) being estimated. Thus a smaller value of \( SSR \) close to zero indicates a better fit. Those values \( \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_m \) that minimise \( SSR \) result in the best fit and thus the estimators are obtained. This method is denoted by PDF in the following text.

Another option is fitting the cumulative distribution function of a specified distribution to the empirical cumulative distribution function of an image. Thus the problem is similarly reduced to minimising the summed square of residuals

\[
SSR = \sum_{x=0}^{1} |F_S(x) - F_X(x; \theta_1, \theta_2, \ldots, \theta_k)|^2
\] (3.42)

where

\[
F_S(x) = \sum_{0 \leq t \leq x} p_s(t)
\] (3.43)

is the empirical cumulative distribution function obtained from the image and

\[
F_X(x; \theta_1, \theta_2, \ldots, \theta_k) = \sum_{0 \leq t \leq x} p_X(t; \theta_1, \theta_2, \ldots, \theta_k)
\] (3.44)

is the cumulative distribution function of the hypothesised distribution model. This method is denoted by CDF in the following text.

Parameter estimation by these three methods described above can be generally re-
garded as unconstrained nonlinear optimisation problems and the simplex search method of Nelder and Mead [58,59] is suitable for these nonlinear problems, in which only function evaluations are involved and an initial estimate of the variables is required.

3.2.3 The Method of Moments and Initial Estimation of Parameters

All of the three methods shown above for parameter estimation require some sensible initial values for the parameters. These initial values can be derived from simulated images by the method of moments [56, 57]. The underlying idea of this method is to equate sample statistics, such as the sample mean and variance, to the corresponding population moments.

If there are $k$ unknown parameters $\theta_1, \theta_2, ..., \theta_k$ to be estimated, then $k$ simultaneous equations need to be established in which the $\nu$-th sample moment equates to the corresponding $\nu$-th population moment. Solving these equations for the unknown parameters yields the moment estimators $\hat{\theta}_1, \hat{\theta}_2, ... \hat{\theta}_k$.

The $\nu$-th sample moment $m_{\nu}$ is given by

$$m_{\nu} = \frac{1}{N} \sum_{i=1}^{N} x_i^\nu$$}

where the $x_i$’s are pixel values of a simulated B-scan image which has $N$ pixels in total.

The Rayleigh Distribution

For the Rayleigh distribution, there is only one parameter $\sigma$ to be estimated. Since both the first moment (population mean) and the second moment (population variance) have closed form expression, either of them can be used to solve for $\sigma$. When the first moment is used, from Equation (3.5) we have

$$E[X] = \sqrt{\frac{\pi}{2}} \sigma = m_1 = \bar{x}$$

and the estimator is

$$\hat{\sigma} = \sqrt{\frac{2}{\pi}} \bar{x}.$$  

For the second moment, from Equation (3.6) we have

$$E[X^2] = 2\sigma^2 = m_2 = \bar{x}^2$$

and the estimator is

$$\hat{\sigma} = \sqrt{\frac{\bar{x}^2}{2}}.$$
3.2. PARAMETER ESTIMATION

The Rician Distribution

For the Rician distribution there are two parameters $\sigma$ and $s$, to be estimated. From Equation (3.10) and (3.11) we have

\[
E[X^2] = 2\sigma^2 + s^2 = m_2
\]  
(3.50)

and

\[
E[X^4] = 8\sigma^4 + 8\sigma^2s^2 + s^4 = m_4
\]  
(3.51)

which yield the following estimators

\[
\hat{s} = \sqrt[4]{2m_2^2 - m_4}
\]  
(3.52)

and

\[
\hat{\sigma} = \sqrt{\frac{m_2 - \hat{s}^2}{2}}.
\]  
(3.53)

The K Distribution

For the K distribution there are two parameters $\alpha$ and $b$, to be estimated. Usually the second moment is used and the other one can either be the fourth moment or the sixth moment. When the second and the fourth moment are used, from Equation (3.14) we have

\[
\frac{E[X^4]}{(E[X^2])^2} = 2 \left(1 + \frac{1}{\alpha}\right) = \frac{m_4}{m_2^2}
\]  
(3.54)

thus

\[
\hat{\alpha} = \frac{m_4}{m_2^2} - 2
\]  
(3.55)

When the second and the sixth moment are used, from Equation (3.15) we have

\[
\frac{E[X^6]}{(E[X^2])^3} = 6 \left(1 + \frac{1}{\alpha}\right) \left(1 + \frac{2}{\alpha}\right) = \frac{m_6}{m_2^3},
\]  
(3.56)

thus

\[
\hat{\alpha} = \frac{9 + \sqrt{9 + \frac{12m_6}{m_2^2}}}{\frac{m_6}{m_2^2} - 6}
\]  
(3.57)

and considering Equation (3.13) $\hat{b}$ can be derived from $\hat{\alpha}$ as

\[
\hat{b} = 2\sqrt{\frac{\hat{\alpha}}{m_2}}.
\]  
(3.58)
The Fisher-Tippett Distribution

For the Fisher-Tippett distribution, there are two parameters, \( \alpha \) and \( \beta \) to be estimated. From Equation (3.20) we have

\[
E[(X - E[X])^2] = \frac{(\pi \beta)^2}{6} = m_2 - m_1^2, \tag{3.59}
\]

thus

\[
\hat{\beta} = \frac{\sqrt{6(m_2 - m_1^2)}}{\pi}. \tag{3.60}
\]

Considering Equation (3.19), the estimator of \( \alpha \) is

\[
\hat{\alpha} = m_1 + \gamma \hat{\beta}. \tag{3.61}
\]

The Nakagami Distribution

For the Nakagami distribution there are two parameters, \( m \) and \( \Omega \), to be estimated. From Equation (3.23) we have

\[
\hat{\Omega} = m_2 \tag{3.62}
\]

and from Equation (3.25) we have

\[
\hat{m} = \frac{m_2^2}{m_4 - m_2^2}. \tag{3.63}
\]

The Weibull Distribution

Since \( f(\beta) \) is a monotonically decreasing function of \( \beta \), a look-up table (LUT) pairing pre-computed values of \( f(\beta) \) and \( \beta \) can be established beforehand, as suggested in [51]. From Equation (3.30) we have

\[
f(\beta) = \frac{m_2}{m_1^2} \tag{3.64}
\]

and then estimator \( \hat{\beta} \) can be obtained from the LUT. After \( \hat{\beta} \) is obtained, from Equation (3.28) we have the estimator for \( \alpha \)

\[
\hat{\alpha} = \frac{m_1}{\Gamma \left(1 + \frac{1}{\beta}\right)}. \tag{3.65}
\]

The Generalised Gamma Distribution

For the generalised Gamma distribution, estimation of the parameters is based on the logarithm of the data and the corresponding moments [51,60]. Specifically, the following
3.3. GOODNESS-OF-FIT TEST

The Generalised Fisher-Tippett Distribution

For the generalised Fisher-Tippett distribution, \( c \) is initially set to 1.0. \( \beta \) is estimated as in Equation (3.60). For parameter \( \alpha \), from Equation (3.36) we have

\[
E[X] = \gamma \beta + 1 - \alpha = m_1
\]

thus

\[
\hat{\alpha} = \gamma \hat{\beta} + 1 - m_1. 
\]

After these moment estimators are obtained, they are presented as the initial values to the method of MLE, which yields a final solution.

3.3 Goodness-of-Fit Test

Up until now we have been concerned about what the parameters would be if the sample data statistically followed a specific family/type of distributions. After parameter estimation using the method of MLE, it is necessary to determine if the sampled population “really” follows the hypothesised distribution, in other words, to what extent the
A postulated probability model agrees with the sample data.

Concretely we would like to know how well the observed frequencies of occurrence of observations fit the expected frequencies obtained from the hypothesised theoretical distribution specified by the parameters. By comparing the observed frequencies with the corresponding expected frequencies, we have to infer whether their discrepancies are trivial and likely to occur due to sampling variability, or are significant indicating that the hypothesised distribution is not reasonable. Accordingly such a test is referred to as a goodness-of-fit test [49, 56, 57, 61, 62].

There are many possible procedures for testing goodness-of-fit. Among them the two most frequently used are the chi-squared ($\chi^2$) goodness-of-fit test and the Kolmogorov-Smirnov (K-S) one-sample test. The chi-squared goodness-of-fit test was employed in [34, 42, 45, 63] to test the hypotheses that envelope of backscattered signal follows the specified distributions and the K-S test was used in [37, 51].

### 3.3.1 The Chi-Squared Goodness-of-Fit Test

To perform a chi-squared goodness-of-fit test, the first step is to group the observations (that is, the pixels of a B-scan image in the following text) into $n$ mutually exclusive and exhaustive classes (also referred to as “cells” or “bins”). The number of observations (pixel counts) falling into each class is the observed frequency of that class, denoted by $O_i$, $\sum_{i=1}^{n} O_i = N$, where $N$ is the total number of observations.

For each class there is a corresponding expected frequency $E_i$, which can be obtained from the hypothesised probability function as $E_i = Np_i$, where $p_i$ is the probability a random observation belongs to the class $i$ when the null hypothesis is true. The null hypothesis $H_0$ and the alternative hypothesis $H_A$ can be respectively stated by:

- $H_0$: The sample data have been drawn from a population that follows a preconceived distribution. In other words, there is no discrepancy between the observed frequency and the expected frequency: $O_i = E_i$, $1 \leq i \leq n$.

- $H_A$: The sample data do not come from a population that follows the specified distribution.

The test statistic is given by

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

(3.72)

which approximately follows a $\chi^2$ distribution with $n - k - 1$ degrees of freedom when the null hypothesis is true. $k$ is the number of parameters of the hypothesised distribution that have to be estimated from the sample data. If the hypothesised distribution fits very well to the sample data, the discrepancies between the observed frequency and the expected frequency will be trivial, resulting in a small $\chi^2$-value. In contrast, a large $\chi^2$-value indicates a poor fit and may lead to the rejection of $H_0$. Thus the critical region
falls in the right tail of the $\chi^2$ distribution.

The decision rule is to reject $H_0$ if $\chi^2 > \chi^2_{\alpha,n-k-1}$, where $\chi^2_{\alpha,n-k-1}$ is the critical value at the $\alpha$ level of significance. Note for the approximation of the test statistic to the $\chi^2$ distribution to be adequate, $N$ should be no less than 30 and every expected frequency $E_i$ should be no less than five. Any $E_i \leq 5$ has to be combined with the adjacent classes until this minimum frequency requirement has been met. Accordingly the number of degrees of freedom of the critical value must be reduced to accommodate to the final actual number of classes [49, 57, 61, 62].

### 3.3.2 The Kolmogorov-Smirnov One-Sample Test

The chi-squared test was designed for use with categorical data. If samples are drawn from continuous populations it requires grouping of data, and thus does not make complete use of the data available. The Kolmogorov-Smirnov (K-S) test [61, 62] was specifically designed for use with continuous data. It allows us to examine the goodness-of-fit for each of the $N$ observations, rather than $n$ classes as in the chi-squared test. While the chi-squared test is based on discrepancies between the observed and expected frequency (the histograms), the K-S test is focused on the differences between the hypothesised cumulative distribution function $F_0(x)$ and the empirical cumulative distribution function $S(x)$.

The null hypothesis $H_0$ and the alternative hypothesis $H_A$ for a two-sided K-S test can be respectively summarised as:

- $H_0 : F(x) = F_0(x)$, where $F(x)$ is the cumulative distribution function of the underlying population represented by the empirical cumulative distribution function $S(x)$ observed from the sample data. $F_0(x)$ is the cumulative distribution function of the hypothesised distribution the population is expected to follow.
- $H_A : F(x) \neq F_0(x)$.

The test statistic, called the Kolmogorov-Smirnov one-sample statistic, is defined by

$$D_{KS} = \sup_x |S(x) - F_0(x)|$$

$$= \max_x (|S(x) - F_0(x)|, |S(x - \varepsilon) - F_0(x)|)$$

(3.73)

where $\sup_x$ means the supremum over all $x$. $\varepsilon$ denotes any small positive number. In practice, the above equation becomes

$$D_{KS} = \max_{1 \leq i \leq n} [\max(|S(x_i) - F_0(x_i)|, |S(x_{i-1}) - F_0(x_i)|)]$$

(3.74)

---

According to [61, 62] five is a conservative value, though. It was suggested that $E_i$ should be no less than 1 [61] or 1.5 [62] in most practical applications. Nevertheless, we are on the conservative side in the following text.
where \( n \) is the number of distinct values of \( x \) and \( S(x_0) = 0 \).

Unlike the chi-squared test statistic, which depends on the hypothesised distribution and only approaches a \( \chi^2 \) distribution, the K-S one-sample test statistic \( D_{KS} \) does not rely on \( F_0(x) \) as long as \( F_0(x) \) is continuous, and so is a distribution-free statistic. Graphically, this statistic denotes the greatest vertical distance between \( F_0(x) \) and \( S(x) \). If the underlying population from which the sample data have been drawn really follows the hypothesised distribution, the curve of \( S(x) \) should agree very well with that of \( F_0(x) \) and \( D_{KS} \) should be fairly small. On the contrary, if \( H_0 \) is not true, large deviations of \( S(x) \) from \( F_0(x) \) and hence a large \( D_{KS} \) are expected.

Note that although it is referred to as “two-sided”, the test statistic is upper-tailed only. The decision rule for a two-sided K-S test at the \( \alpha \) level of significance is to reject \( H_0 \) if \( D_{KS} > D_{N,\alpha} \) where \( D_{N,\alpha} \) is the critical value at the \( \alpha \) level of significance and approximately

\[
D_{N,\alpha} = \frac{d_\alpha}{\sqrt{N}} \tag{3.75}
\]

when the number of observations \( N > 40 \). \( d_\alpha \) for various levels of significance \( \alpha \) can be found in the Appendix of [61, 62]. For \( \alpha = 0.01 \), \( d_{0.01} = 1.63 \).

### 3.4 Summary

In this chapter we have laid the foundation for analysis of statistical properties of simulated ultrasound B-scan images, to be presented in the next chapter. In section 3.1 we reviewed the speckle distribution models. Especially, we empirically proposed the generalised Fisher-Tippett distribution, which will be applied to analyse the real ultrasound images in section 4.3.

In section 3.2 we presented the methods for parameter estimation, including the method of MLE, PDF, CDF and the method of moments. The method of moments is mainly used to generate an initial estimation of the parameters, which is fed into the other three methods for a final solution.

While the methods for parameter estimation tell us what the parameters would be if the sample data statistically followed a specific distribution, the subsequent goodness-of-fit tests verifies whether the sampled population “really” follows the hypothesised distribution. Two goodness-of-fit tests were introduced in section 3.3, namely the chi-squared test and the Kolmogorov-Smirnov one-sample test. Having described the speckle distribution models, methods for parameter estimation, and goodness-of-fit tests, we are ready to analyse statistical properties of simulated ultrasound B-scan images, which is the topic of next chapter.
Chapter 4

Computer Simulation of B-mode Ultrasound Imaging

In this chapter a computer simulation of B-mode ultrasound imaging is presented in an attempt to gain a better understanding of the nature of ultrasonic pulse-echo imaging. Several questions will be addressed by the computer simulation, such as how the B-scan images are formed from the backscattered pulse-echo signal, how the speckle is related to the image formation process, and whether the speckle is solely a characteristic of the imaging system or a result of interaction between the tissue and the imaging system. The computer simulation is expected to provide us with an insight into the appearance and statistical properties of B-scan ultrasound images, which is essential to the design of the speckle reduction techniques later.

In section 4.1 a hypothetical B-mode ultrasonic imaging system will be described, the features of which include system (transducer) modelling and tissue modelling, envelope detection, logarithmic compression and simulation of scatterer density. Section 4.2 presents simulated B-scan images for various scenarios and the analysis of their statistical properties, which inspired the statistical analysis of the ultrasound images of the prostate, described in section 4.3.

4.1 Hypothetical B-mode Ultrasonic Imaging System

Figure 4.1 shows the schematic setup of a hypothetical B-mode ultrasonic imaging system. \( D \) represents a cross section of the background medium and three discs \( A \), \( B \) and \( C \) represent cross sections of objects of interest. Their acoustic impedances are different as depicted by the different shades of the discs. \( x \) represents the direction of insonification (or the axial direction). A simple linear B-mode scan is assumed and the transducer moves in a direction perpendicular to the sound beam. This direction is referred to as the lateral or azimuthal direction and denoted by \( y \) in Figure 4.1.

We have the following assumptions for the hypothetical ultrasonic imaging system: zero attenuation, no multiple scattering, maximum transducer bandwidth, and constant...
4.1 HYPOTHETICAL B-MODE ULTRASONIC IMAGING SYSTEM

Figure 4.1: A schematic setup of a hypothetical B-mode ultrasonic imaging system. \( x \) represents the direction of insonification and \( y \) the direction of transducer movement. \( D \) represents a cross section of the background medium and \( A, B \) and \( C \) are cross sections of three objects of different acoustic impedance.

velocity of sound through the medium or biological tissue\(^1\). Zero attenuation means that the time-gain-control (TGC) perfectly compensates for the attenuation of the echo signal returned by scatterers in the depth. It is assumed that the backscattered signal from each scatterer directly returns to the transducer, that is, there is no interaction between the backscattered signal from different scatterers.

Computer simulation of ultrasonic B-scanning involves modelling the imaging system (especially the transducer), the tissue, envelope detection, logarithmic compression and simulation of scatterer density.

4.1.1 Imaging System Modelling

An ultrasound B-mode image \( A(x, y) \) formed by an imaging system is determined by (Appendix 1 of [64])

\[
A(x, y) = H(x, y) * T(x, y)
\]  

\( ^1 \)In the following text the term “medium” and “tissue” are equivalent.
where * denotes a convolution process. $H(x, y)$ represents the point spread function (PSF) of the imaging system, which is assumed to be linear and space invariant. $T(x, y)$ is the acoustic impulse response of the medium, which depends on its density and acoustic compressibility.

The PSF $H(x, y)$ is not circularly symmetric, but assumed to be separable and characterised by two component functions $H_1(x)$ and $H_2(y)$, that is, $H(x, y) = H_1(x)H_2(y)$. $H_1(x)$ represents the axial ultrasound pulse emitted by the transducer and $H_2(y)$ characterises the lateral beam profile of the transducer [64].

$H_1(x)$ can be defined by a Gaussian-weighted sinusoid of limited duration (a Gabor function)

$$H_1(x) = \sin \left( \frac{2\pi f_0 c}{c} x \right) e^{-x^2/2\sigma_x^2}$$

(4.2)

where $f_0$ is the centre (carrier) frequency of the transducer, $c$ is the sound speed through the medium, and $\sigma_x$ characterises the ultrasound pulse width. This is similar to the transmitted pulse of the transducer model of the computer simulation in [26, 33, 36, 42, 44, 65].

The lateral beam profile of the transducer $H_2(y)$ is assumed to be a Gaussian function (that is, no sidelobes$^2$)

$$H_2(y) = e^{-y^2/2\sigma_y^2}$$

(4.3)

where $\sigma_y$ characterises the beam width of the transducer. Both the pulse shape and beam profile are assumed invariant through the medium in the following text.

Thus the image is generated by consecutively convolving the tissue model $T(x, y)$ with the two components of the PSF, $H_1(x)$ and $H_2(y)$

$$A(x, y) = H_2(y) * H_1(x) * T(x, y).$$

(4.4)

### 4.1.2 Modelling the Tissue

Apart from the imaging system, the formation of a B-mode ultrasound image also relies on the acoustic impulse response of a given medium. Unfortunately it is very difficult to precisely determine the acoustic impulse response as it requires knowledge of characteristics of the medium on a scale far less than the ultrasound wavelength [66]. As alternatives two stochastic models are proposed, namely, the inhomogeneous continuum model [64, 67, 68] and the discrete scatterer model [66].

In the inhomogeneous continuum model the target tissue is considered to be made up of a continuous distribution of elemental scattering sources whose density and compressibility fluctuate from place to place about their mean values [68]. For simplicity of analysis, usually the density is assumed to be constant. The compressibility $\beta(x, y)$ of the tissue is modelled as a random function of position whose auto-correlation function is a Gaussian

$^2$A beam profile with sidelobes can be approximated by a sinc function appropriate to a transducer with a rectangular aperture [24, 66].
4.1. HYPOTHETICAL B-MODE ULTRASONIC IMAGING SYSTEM

Function

\[ R(\beta(x, y)) = e^{-(x^2 + y^2)/a^2} \] (4.5)

where \( a \) is the correlation length. \( \beta(x, y) \) can then be synthesised from the auto-correlation function \( R(\beta(x, y)) \) so that its power spectrum is the Fourier transform of \( R(\beta(x, y)) \) and its phase spectrum is a matrix of random numbers. Then the impulse response of the tissue is

\[ T(x, y) = \nabla^2 \beta(x, y) \] (4.6)

where \( \nabla^2 \) denotes the Laplacian operator.

In the discrete scatterer model the target tissue is regarded as a collection of point scatterers at random positions in a homogeneous region. It can be shown that the two models are equivalent if the width of the auto-correlation function \( R(\beta(x, y)) \) is small compared to the ultrasonic wavelength and pulse width [66].

In the following simulation the discrete scatterer model is assumed and the tissue is modelled as

\[ T(x, y) = T_{ph}(x, y) + T_{ph}(x, y) \cdot n(x, y) \] (4.7)

where \( T_{ph}(x, y) \) is an ideal phantom of the target tissue, which is an ensemble of a number of homogeneous regions of different acoustic impedance appropriate to the underlying tissue. \( n(x, y) \) is a random field of normal distribution with mean zero and variance one.

4.1.3 Envelope Detection

Since the analytical expressions of the Fourier transform of \( H_1(x) \) and \( H_2(y) \) are readily available, the convolutions in Equation (4.4) can be accomplished by multiplications in the frequency domain via the fast Fourier transform (FFT). However, the “image” obtained from Equation (4.4) is actually a series of radio frequency (RF) signals (known as A-mode lines). One final step remains before presentation to the B-mode visual display, that is, demodulation or envelope detection.

Demodulation or envelope detection can be accomplished by the analytic signal\(^3\) approach [69–72]. Since the observed backscatter signal \( A(x, y) \) is a real and bandpass signal, the associated analytic signal \( A_a(x, y) \) can be constructed from \( A(x, y) \) by

\[ A_a(x, y) = A(x, y) + j \hat{A}(x, y) \] (4.8)

where \( j = \sqrt{-1} \) and \( \hat{A}(x, y) \) is the Hilbert transform (or quadrature function) of \( A(x, y) \). \( \hat{A}(x, y) \) is actually \( A(x, y) \) shifted by \( \pi/2 \) in phase. Then the envelope function \( I(x, y) \) of \( A(x, y) \) can be computed as

\[ I(x, y) = |A_a(x, y)| = \sqrt{A^2(x, y) + \hat{A}^2(x, y)} \] (4.9)

\(^3\)An analytic signal is a complex function whose Fourier transform vanishes for negative frequencies. In other words, a signal is “analytic” if it is one-sided in the dual Fourier domain.
A quick way to obtain $I(x, y)$ is to make use of the Fourier transform of $A(x, y)$. $A(x, y)$ is first transformed to the frequency domain, followed by multiplication by a step function in the frequency domain to make all values at negative frequency vanish. Finally the inverse Fourier transform gives the analytic signal $A_a(x, y)$. This is summarised as

$$A_a(x, y) = \mathcal{F}^{-1} \left[ \left( 1 + \frac{\omega}{|\omega|} \right) \mathcal{F}(A(x, y)) \right]$$

where $\omega$ is the radian frequency. In commercial ultrasound imaging systems, however, the envelope is usually obtained from the RF signal by half-wave or full-wave rectification or square law detection followed by low pass filtering, which is a fairly accurate approximation to that of the analytic signal approach [29, 64].

### 4.1.4 Logarithmic Compression

After envelope detection, the signal is ready for presentation to the display device. In clinical ultrasound imaging systems, however, the envelope detected signal is subjected to further nonlinear amplification to reduce the dynamic range of signal amplitudes, which is usually much larger than what the display device can handle.

Usually logarithmic compression is applied, which amplifies low level signals to a greater degree than high level signals. The logarithmic amplification process can be modelled as [27, 38–40, 54]

$$I_Y = N_d \ln(I_X) + N_g$$

where $I_X$ is the signal before amplification and $I_Y$ is the signal after amplification. $N_d$ is related to the dynamic range of the system at the stage of RF signal (before B-mode image presentation). $N_g$ characterises the overall gain setting of the amplifier. This model will be referred to as “standard” logarithmic compression in the following text.

However, the signal amplified by this model tends to have a negative (left) skewed probability density function (p.d.f.) [27, 38]. This does not agree with the histograms of our ultrasound images of the prostate produced by the Proscan® Plus imaging system, which are strongly positive (right) skewed (see Figure 4.19-4.22). This implies that a different signal compression curve might be used.

To be consistent with our real ultrasound images of the prostate, the following signal compression model is adopted in our computer simulation

$$I_Y = \frac{\ln(I_X + 1)}{\ln(I_{X_{\text{max}}} + 1)}$$

where $I_{X_{\text{max}}} = \max(I_X)$ is the maximum of the input signal. This model still keeps the right skewness of the original input signal in the range $0 \leq I_x \leq 1$. It is referred to as *ad hoc* logarithmic compression in the following text.
4.1. HYPOTHETICAL B-MODE ULTRASONIC IMAGING SYSTEM

Figure 4.2: Simulated B-scan image of a homogeneous area by the hypothetical B-mode ultrasonic imaging system. The relevant simulation parameters are $\sigma_x = 1$, $\sigma_y = 2.5$. Scatterer density $d_s = 1.0$ and no logarithmic amplification is applied. (a) Simulated image of a homogeneous area, image size 256 $\times$ 256 pixels. (b) After histogram equalisation.

Figure 4.3: Comparison of a horizontal and a vertical line profile through the simulated B-scan image shown in Figure 4.2. (a) A horizontal line profile through the simulated image. (b) A vertical line profile. Note the noise in the vertical profile fluctuates at a much higher rate than that in the horizontal profile.
4.1.5 Scatterer Density

The scatterer density is simulated by a coefficient $d_s (0 < d_s \leq 1.0)$, which is related to the number of scatterers in a resolution cell of the imaging system when it produces fully-developed speckles ($d_s = 1.0$). For example, if the ultrasound pulse parameter $\sigma_x = 1$ and the beam width parameter $\sigma_y = 2.5$, then the number of scatterers per resolution cell is approximately $90$ when $d_s = 1.0$. The scatterer density coefficient $d_s = 0.5$ will result in approximately 45 scatterers per resolution cell.

Simulated B-scan images will be presented next in section 4.2. After the simulated images are obtained, we need to analyse the statistical properties of simulated B-scan images and verify whether the pixel values of the simulated images follow various hypothesised distribution models for different situations. For example, we would like to check whether the pixel values of a simulated B-scan image of a homogeneous phantom follow the Rayleigh distribution. This involves parameter estimation and goodness-of-fit tests for various probability distribution models.

4.2 Simulated Images and Statistical Analysis

In this section simulated B-scan images produced by the hypothetical ultrasonic imaging system are presented. Various distribution models are fitted to the images in an attempt to gain an insight into the statistical properties of speckles. The method of MLE, CDF and PDF are used to estimate parameters of a hypothesised distribution, followed by chi-squared and K-S goodness-of-fit tests to verify whether the hypothesised distribution is reasonable.

4.2.1 Homogeneous Phantoms

Figure 4.2(a) shows the simulated B-mode image of a homogeneous phantom of $256 \times 256$ pixels. The relevant simulation parameters are $\sigma_x = 1$, $\sigma_y = 2.5$. Scatterer density is assumed to be $d_s = 1.0$ and there is no logarithmic amplification applied. Figure 4.2(b) shows the image after histogram equalisation. The SNR of the image is 1.918.

As we can see, the simulated image is dominated by the distracting granular appearance, which is immediately apparent even when we have a glance at the image. This kind of well developed speckle pattern is a familiar sight in the ultrasound images of the prostate, as we have seen in the previous chapters, in which the exhibited “texture” does not necessarily represent separately resolvable structures in the tissue. This is now confirmed by our simulation scan of the homogeneous phantom.

Figure 4.3(a) and (b) respectively show a horizontal (perpendicular to the insonification beam) and a vertical (parallel to the insonification beam) line profile randomly selected

---

4This can be implemented by simulating $n(x, y)$ in Equation (4.7) as a sparse normally distributed random matrix using MATLAB’s function `sprandn`.

5Recall that the two component functions (Equation (4.2)-(4.3)) of the PSF are of Gaussian shape, then the size of a resolution cell is approximately $6\sigma_x \times 6\sigma_y = 36\sigma_x\sigma_y$. 

59
from the simulated image shown in Figure 4.2. It is immediately apparent that the noise in the vertical profile fluctuates at a much higher rate than that of the horizontal profile. This is attributed to the different axial and lateral resolution of the imaging system.

Nevertheless, the speckle noise in both directions is an undesired artifact which masks off the real signal and make it very difficult to visually assess the ultrasound images. Therefore a suitable image processing routine is needed to distinguish the real signal from the fluctuation due to speckle and reliably suppress it.

**Statistical Properties of the Simulated Images**

Since the scatterer density \( d_s = 1.0 \), with transducer parameters \( \sigma_x = 1 \) and \( \sigma_y = 2.5 \), for this case the number of scatterers in a resolution cell of the imaging system is large enough to justify the applicability of the *Central Limit Theorem*. Therefore fully formed speckle is expected and the amplitude of the speckle should follow a Rayleigh distribution. To verify this hypothesis, the p.d.f. of the Rayleigh distribution is fitted to the histogram of the image and the parameter \( \sigma \) is estimated (thus number of parameters \( k = 1 \)) by the method of MLE, followed by the chi-squared goodness-of-fit test and K-S test.

Figure 4.4 compares the histogram (light-colour solid line) of the simulated image in Figure 4.2(a) with the fitted p.d.f. (dark-colour dashed line) of a Rayleigh distribution (\( \hat{\sigma} \approx 0.21 \)). As expected, the Rayleigh distribution with parameters estimated agrees very well with the fitted p.d.f. of the Rayleigh distribution.
well with the histogram of the image.

A chi-squared goodness-of-fit test was conducted at the $\alpha = 0.05$ level of significance and the pixels were grouped into 256 bins based on their gray scale values. The test statistic is $\chi^2 = 144.852 < \chi^2_{0.05,221} = 256.68$. Thus it provides strong evidence that the null hypothesis that the pixel gray scale values of the image follow a Rayleigh distribution should not be rejected. Note that the degrees of freedom of the chi-squared critical value is 221, which is less than 254 ($256 - k - 1$). This is due to adjustment of number of bins to satisfy the restriction of minimum frequency of 5 for each bin. A K-S test gives the result $D = 0.003578$ with a $p$-value 0.3711. The critical value at the $\alpha = 0.01$ level of significance is $D_{0.01,256^2} = 1.63/256 = 0.0064$. Thus the K-S test statistic also supports that the null hypothesis should not be rejected.

Rather than jumping to conclusions on the basis of only one image, 500 B-scan images of homogeneous areas were simulated and analysed. The Rayleigh distribution was fitted to each image by the method of MLE, CDF and PDF, followed by goodness-of-fit test, a chi-squared test at the $\alpha = 0.05$ level of significance and a K-S test at the $\alpha = 0.01$ level of significance. The results of the goodness-of-fit tests are illustrated in Figure 4.5 and Figure 4.6.

**Results of the Chi-Square Goodness-of-Fit Test**

Figure 4.5 compares the chi-squared test statistics (denoted by $\chi^2$ in the figure) of the three parameter estimation methods, MLE, CDF and PDF against the corresponding critical values (denoted by $\chi^2_{CV}$ in the figure). The diagonal line represents $\chi^2 = \chi^2_{CV}$. As we can see, the overwhelming majority of the test results fall below the $\chi^2 = \chi^2_{CV}$ line, that is, in the $\chi^2 < \chi^2_{CV}$ region, strongly supporting that the null hypothesis should not be rejected.

The method of MLE and CDF both have two cases (the two stars lying close to the $\chi^2 = \chi^2_{CV}$ line) that the null hypothesis is nearly rejected but a subsequent comparison shows that the test statistics are still marginally outside the rejection region. The results of the method of PDF have two rejections (the two dots above the $\chi^2 = \chi^2_{CV}$ line), which are fairly insignificant compared to the large number of tests. Note that in the figure the large number of cases where the cross, plus sign and the dot are overlapping, which is a strong indication that the three methods of parameter estimation are comparable in the sense of chi-squared goodness-of-fit, although it seems that the method of MLE and CDF are marginally “better”. We will test this hypothesis later.

**Results of the Kolmogorov-Smirnov Goodness-of-Fit Test**

The corresponding K-S test statistics for the three methods are illustrated in Figure 4.6. The horizontal line across the figure represents the critical value of the K-S test at $\alpha = 0.01$ level of significance for $256 \times 256$ observations, which is 0.0064. As we can see the number of points above the line of critical value (which indicate the rejections of null hypothesis) are rather sparse compared to the overwhelming majority below the line. Actually the
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

Figure 4.5: Results of the chi-squared goodness-of-fit test on 500 simulated B-scan images of homogeneous areas (see text for details).

Figure 4.6: Results of the Kolmogorov-Smirnov goodness-of-fit test on 500 simulated B-scan images of homogeneous areas (see text for details).
method of CDF has 10 rejections, followed by the method of PDF, 24 rejections and the method of MLE 48 rejections.

The numbers of rejections of the three methods are slightly higher than the expected number for an experiment of 500 tests. However, as we can see, when using the method of CDF, 98% of the tests lie below the critical line. After the overwhelming support from the chi-squared test being taken into account, it is safe to say that the null hypothesis should not be rejected.

The method of CDF has the least number of rejections among that of the three methods. This leads us to hypothesise that the method of CDF is the best among the three methods for parameter estimation in the sense of K-S goodness-of-fit. We will formally compare the three methods later.

The SNRs of the 500 images were also recorded. A $t$-test was performed at significance level $\alpha = 0.001$ to test if the null hypothesis $\mu_{SNR} = 1.9131$ is true. The result is that the null hypothesis could not be rejected (with a $p$-value 0.9723). A 99.9% confidence interval (C.I.) on the mean is $[1.9039, 1.9160]$, which brackets the theoretical (and hypothesised) value of 1.9131 for the Rayleigh distribution model.

In summary, it is apparent from the above analysis that if the number of scatterers in a resolution cell of the imaging system is large and no logarithmic amplification is applied, the pixel values of pulse-echo B-scan images follow the Rayleigh distribution. On the other hand, it also confirms that our computer simulation is reasonable.

**Comparison of the Three Methods for Parameter Estimation**

Now the three methods for parameter estimation, MLE, CDF and PDF, are formally compared in the sense of chi-squared and K-S goodness-of-fit. We first compare the chi-squared goodness-of-fit test statistics.

Figure 4.7(a) compares the boxplots of chi-squared goodness-of-fit test statistics of the three methods. The lower and the upper sides of the rectangles respectively correspond to the lower and upper quartiles of the data. The horizontal lines across the rectangles represents the medians. The whiskers extend to $1.5IQR$ from either end of the rectangles. The means are denoted by “$\times$” and outliers by “$+$”. As we can see, the boxplots look quite similar, which is a strong indication that their chi-squared test statistics are not significantly different.

Next we formally test this hypothesis, that is, we consider the following null hypothesis and alternative hypothesis:

$H_0$: The method of MLE, CDF and PDF are equivalent in parameter estimation (for the simulated images of homogeneous phantoms) in the sense of chi-squared goodness-of-fit, that is, the medians of their chi-squared test statistics are not significantly different, $\tilde{\chi}^2_{MLE} = \tilde{\chi}^2_{CDF} = \tilde{\chi}^2_{PDF}$.

$H_A$: The medians of chi-squared goodness-of-fit test statistics of the three methods are significantly different.
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

Figure 4.7: Comparison of the boxplots of the goodness-of-fit test statistics of the method of MLE, CDF and PDF. (a) Their chi-squared test statistics are not significantly different. (b) The difference between their K-S test statistics are significant. The method of CDF provides the best fit.

<table>
<thead>
<tr>
<th>NULL HYPOTHESIS $H_0$</th>
<th>$T_W$</th>
<th>$p$-value</th>
<th>CRITICAL VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2_{MLE} = \chi^2_{CDF}$</td>
<td>-0.5537</td>
<td>0.5798</td>
<td>$Z_{0.05} = -1.96$</td>
</tr>
<tr>
<td>$\chi^2_{MLE} = \chi^2_{PDF}$</td>
<td>-1.6341</td>
<td>0.1022</td>
<td></td>
</tr>
<tr>
<td>$\chi^2_{CDF} = \chi^2_{PDF}$</td>
<td>-1.0942</td>
<td>0.2738</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Results of the Wilcoxon-Mann-Whitney test on the chi-squared goodness-of-fit test statistics of the method of MLE, CDF and PDF. The null hypotheses are not rejected.

<table>
<thead>
<tr>
<th>NULL HYPOTHESIS $H_0$</th>
<th>$T_W$</th>
<th>$p$-value</th>
<th>CRITICAL VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{MLE} = D_{CDF}$</td>
<td>7.1288</td>
<td>$1.01 \times 10^{-12}$</td>
<td>$Z_{1-0.05} = 1.64$</td>
</tr>
<tr>
<td>$D_{MLE} = D_{PDF}$</td>
<td>3.7022</td>
<td>0.00021</td>
<td>$Z_{1-0.05} = 1.64$</td>
</tr>
<tr>
<td>$D_{CDF} = D_{PDF}$</td>
<td>-3.6694</td>
<td>0.00024</td>
<td>$Z_{0.05} = -1.64$</td>
</tr>
</tbody>
</table>

Table 4.2: Results of the Wilcoxon-Mann-Whitney test on the K-S goodness-of-fit test statistics of the method of MLE, CDF and PDF. The null hypotheses are rejected.
An application of the Lilliefors test for normality [61] shows that the assumption of normality is not reasonable for the results of the chi-squared goodness-of-fit test of the three methods. Therefore the Wilcoxon-Mann-Whitney rank sum test [49,56,61] is chosen to carry out the test.

The Wilcoxon-Mann-Whitney test is a non-parametric hypothesis test procedure involving overall ranking of the data. If the null hypothesis is true and the sizes of samples are no less than 20, then the normal distribution provides a good approximation to the test statistics $T_W$. The decision rule is to reject $H_0$ if $T_W < Z_{\frac{\alpha}{2}}$ or $T_W > Z_{1-\frac{\alpha}{2}}$ for a two-sided test ($T_W < Z_\alpha$ or $T_W > Z_{1-\alpha}$ for a one-sided test), where $\alpha$ is the desired level of significance. $Z_{\frac{\alpha}{2}}$ and $Z_{1-\frac{\alpha}{2}}$ are respectively the $\frac{\alpha}{2}$ and $(1-\frac{\alpha}{2})$ quantile of the normal distribution.

Three null hypotheses $H_0$, $\tilde{\chi}_{MLE}^2 = \tilde{\chi}_{CDF}^2$, $\tilde{\chi}_{MLE}^2 = \tilde{\chi}_{PDF}^2$ and $\tilde{\chi}_{CDF}^2 = \tilde{\chi}_{PDF}^2$ are tested at the level of significance $\alpha = 0.05$. The results of these two-sided tests are summarised in Table 4.1. As we can see, the test statistics are all greater than the critical value $Z_{0.05}^{\frac{\alpha}{2}}$ and the corresponding $p$-values are all significantly greater than 0.05. Thus they provide strong evidence that the null hypotheses could not be rejected. Therefore we can conclude that the three parameter estimation methods are equivalent in terms of chi-squared goodness-of-fit for these images.

Next we compare the Kolmogorov-Smirnov test statistics of the three methods. Figure 4.7(b) shows the boxplots of their K-S test statistics. As we can see, the difference between the medians of their K-S test statistics, $\tilde{D}_{MLE}$, $\tilde{D}_{CDF}$ and $\tilde{D}_{PDF}$, are not trivial. Therefore we test the following hypotheses using the Wilcoxon-Mann-Whitney test:

- $H_0$: $\tilde{D}_{MLE} = \tilde{D}_{CDF}$, with $H_A$: $\tilde{D}_{MLE} > \tilde{D}_{CDF}$;
- $H_0$: $\tilde{D}_{MLE} = \tilde{D}_{PDF}$, with $H_A$: $\tilde{D}_{MLE} > \tilde{D}_{PDF}$;
- $H_0$: $\tilde{D}_{CDF} = \tilde{D}_{PDF}$, with $H_A$: $\tilde{D}_{CDF} < \tilde{D}_{PDF}$.

These one-sided tests are conducted at the level of significance $\alpha = 0.05$.

The test results are listed in Table 4.2. As we can see, the first two test results are all greater than the critical value $Z_{1-0.05}$, thus fall in the rejection region and the corresponding $p$-values are nearly 0. The third test result is less than the corresponding critical value $Z_{0.05}$, thus belongs to the rejection region. The $p$-values is 0.0024. Therefore all three null hypotheses should be rejected and the alternative hypotheses are accepted. This means that the three parameter estimation methods are significantly different in terms of K-S goodness-of-fit, with the method of CDF providing the best fit. This is fairly understandable since in the method of CDF we are explicitly fitting the cdf and the K-S test is cdf “oriented”, that is, on the basis of the differences between the hypothesised cumulative distribution function and the empirical cumulative distribution function.
4.2. Simulation of Images and Statistical Analysis

4.2.2 Homogeneous Phantoms with Effect of Scatterer Density $d_s$

It is a known fact that the number of scatterers per resolution cell must be large enough for the Central Limit Theorem to hold. Once this condition is not satisfied, that is, the number of scatterers in a resolution cell is sparse, the distribution of the pixel values of B-scan images will skew away from the Rayleigh distribution [27].

As an example, Figure 4.8 shows the histogram (light-colour solid line) of a simulated B-scan image with scatterer density reduced to $d_s = 0.5$ and transducer parameters unchanged ($\sigma_x = 1$, $\sigma_y = 2.5$). Thus there are approximately 45 scatterers per resolution cell. Together with the histogram are the fitted p.d.f.’s of a Rayleigh distribution (dark-colour dashed line) and K-distribution (dark-color dash-dot line). The SNR of the simulated image is $1.51 < 1.9131$, thus the Rician distribution is not applicable.

As we can see, the histogram skews away from the fitted p.d.f. of the Rayleigh distribution to become pre-Rayleigh whereas the K-distribution with parameters estimated provides a reasonable fit to the histogram. This indicates that the K-distribution is more appropriate than the Rayleigh distribution to describe the statistical properties of B-scan images when the number of scatterers per resolution cell is sparse.

Next we formally test our hypothesis, that is, simulated B-scan images follow the K
COMPUTER SIMULATION OF B-MODE ULTRASOUND IMAGING

Table 4.3: Number of rejections of the null hypothesis that simulated B-scan images follow the K distribution when the number of scatterers per resolution cell is sparse.

distribution when the number of scatterers per resolution cell is sparse. B-scan images were simulated with the number of scatterers per resolution cell varying from 80 to 8. For each case, 100 images were simulated and fitted to the K distribution by the methods of MLE, CDF and PDF, followed by the chi-squared and K-S goodness-of-fit tests. (The images were also fitted to the Rayleigh distribution, but the goodness-of-fit tests failed for all cases, as expected.) The numbers of rejections of the null hypothesis are summarised in Table 4.3.

The first column of Table 4.3 lists the number of scatterers per resolution cell. The next three columns are the numbers of rejections of the null hypothesis of the chi-squared goodness-of-fit tests for the three parameter estimation methods. The corresponding results of the K-S tests are listed in the last three columns.

As we see from the table, although the K distribution is not appropriate when the number of scatterers reduces to 10 or less (which is almost impossible in reality), it is a reasonable model for characterising the B-scan images for most of the cases listed in the table. Therefore it is safe to say that when the number of scatterers per resolution cell is sparse (but not extremely sparse), the distribution of pixel values of the B-scan images obeys the K distribution.

Another observation is that once the scatterers becomes sparse, they loss their randomness and the corresponding images become K distributed, even the number of scatterers per resolution cell is still high. Any divergence from this observation warrants further research.

For these K distributed images, the three methods for parameter estimation are compared in terms of goodness-of-fit. In the sense of chi-squared goodness-of-fit, the Wilcoxon-Mann-Whitney tests reveal that the methods of MLE and CDF are comparable although the former is marginally better. The method of PDF provides the “worst” fit. In the sense of K-S goodness-of-fit, the method of CDF is significantly better than the other two methods, as expected. The methods of MLE and PDF are comparable although the former provides a marginally better fit between the two methods.
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

(a) The simulated image. It looks rather “over-exposed”. (b) Its negative (left) skewed histogram with the fitted Fisher-Tippett (F-T) distribution.

4.2.3 Homogeneous Phantoms with Effect of Logarithmic Compression

As we have seen in section 4.1.4, logarithmic compression favors weak signals more than strong signals. Therefore the statistical properties of the original signal are expected to change substantially once logarithmic amplification is applied.

As an example, Figure 4.9(a) shows a simulated image of a homogeneous area with the “standard” logarithmic compression applied. The other system factors, such as transducer parameters and scatterer density, are kept the same as that of the image shown in Figure 4.2(a). As we can see, the image look rather “over-exposed”. The SNR of the image is 8.62, far more than 1.9131 (compare to the SNR of the B-scan ultrasound images of the prostate, which is greater than 1.9131 but less than 3).

The histogram of this image is shown in Figure 4.9(b), together with the fitted Fisher-Tippett (F-T) distribution (dashed line). It is immediately apparent that it is completely different from the one shown in Figure 4.4, since it is rather negative skewed. This is not consistent with the appearance of the histograms of the prostate ultrasound images, which are strongly positive (right) skewed. For this reason we will concentrate on the ad hoc logarithmic compression instead in the following text. Nevertheless, investigation of the effect of the “standard” logarithmic compression and the corresponding Fisher-Tippett distribution has inspired us to propose the generalised Fisher-Tippett distribution and apply it to analyse statistical properties of the B-scan ultrasound images of the prostate, as described in the next section.

With the ad hoc logarithmic compression applied, 500 images of a homogeneous phan-
Figure 4.10: Effect of logarithmic compression: histogram (solid line) of a simulated image with *ad hoc* logarithmic compression applied. Note the histogram skews away from the fitted p.d.f. of a Rayleigh distribution (dashed line) but agrees very well with the fitted p.d.f. of a Rician distribution (dash-dot line).

Tom (scatterer density $d_s = 1.0$) were simulated. The images look rather similar to the image shown in Figure 4.2(a) though marginally brighter, so they are not shown here. The mean SNR of these images is 2.05 with a standard deviation of 0.01. This is greater than 1.9131, casting doubt on the assumption that their pixel values still follow the Rayleigh distribution.

The histogram (light-colour solid line) of one of these images is shown in Figure 4.10, together with the fitted p.d.f. of the Rayleigh distribution (dark-colour dash line) and Rician distribution (dark-colour dash-dot line). Since the SNRs of these simulated images are greater than 1.9131, the K-distribution is not applicable.

As we can see, the histogram still keeps positive skewed but becomes post-Rayleigh since it skews away to the right of the fitted p.d.f. of the Rayleigh distribution. However, it agrees very well with the fitted p.d.f. of the Rician distribution. This indicates that the statistics of B-scan images of a homogeneous phantom follow the Rician distribution rather than the Rayleigh distribution when the *ad hoc* logarithmic compression is applied.

Next we formally test our hypothesis that these images follow the Rician distribution. The Rician distribution is fitted to each image by the methods of MLE, CDF and PDF, followed by the chi-squared goodness-of-fit test at the $\alpha = 0.05$ level of significance and the K-S test at $\alpha = 0.01$. 

69
Figure 4.11 illustrates the chi-squared test results after fitting with the three parameter estimation methods. As we can see, the test statistics all fall in the $\chi^2 < \chi^2_{CV}$ region, strongly supporting that the null hypothesis should not be rejected. The K-S test results are similar except that there is one rejection of the null hypothesis for the method of PDF, as shown in Figure 4.12.

Therefore it is safe to draw the conclusion that after application of the ad hoc logarithmic compression, the distribution of pixel values of the simulated B-scan images obeys a Rician distribution.

For these Rician distributed images, the three methods for parameter estimation are compared in terms of chi-squared and K-S goodness-of-fit. The Wilcoxon-Mann-Whitney tests reveal that the methods of MLE and CDF are comparable in the sense of chi-squared goodness-of-fit, although the former is marginally better. The method of PDF provides the “worst” fit. In the sense of K-S goodness-of-fit, however, the method of CDF provides significantly better fit than the other two methods, as expected. The method of MLE yields the “worst” fit among the three methods.

When number of scatterers are sparse and the ad hoc logarithmic compression is applied, the “best-fit” distribution is either the K-distribution or the Rician distribution, depending on whether the number of scatterers is “sparse enough” or not. It was found that with the ad hoc logarithm compression applied, some of the formerly K-distributed images are “promoted” to follow the Rician distribution when the number of scatterers per resolution cell is approximately no less than 40. Otherwise the images remain K distributed. Any divergence from this observation warrants further research.

4.2.4 Phantoms with Coherent Components

In the previous subsections we have been concentrating on homogeneous phantoms. It is found that the speckle amplitude of a simulated B-scan image for a homogeneous phantom normally follows the Rayleigh distribution. But it skews away to the K-distribution when the number of scatterers per resolution cell is not large enough, or to the Rician distribution when the ad hoc logarithmic compression is applied.

In this subsection simulated images of phantoms with specular or coherent components will be discussed. Apart from scatterer density and the ad hoc logarithmic compression discussed in the previous subsections, there are several other factors that affect the appearance and statistical properties of these images:

- **Background Region of the Phantom:** In reality this can be the acoustic impedance of the “background” tissue that encompasses the object of interest. Intuitively, a dark background results in a dim B-scan image whose histogram is rather right skewed with the mode\(^6\) swinging to the left (that is, to the low gray scale values). The SNR of the image is smaller since the mean gray scale value is lower. On the contrary, a bright background yields a bright scan image with greater SNR.

\(^6\)The value with the highest frequency in a set of data, that is, the maximum of a histogram in practice.
Figure 4.11: Results of the chi-squared goodness-of-fit test on 500 simulated B-scan images of homogeneous phantoms with the ad hoc logarithmic compression applied.

Figure 4.12: Results of the Kolmogorov-Smirnov goodness-of-fit test on 500 simulated B-scan images of homogeneous areas with the ad hoc logarithmic compression applied.
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

- **Contrast of Object/Background Regions:** In reality this may model the difference in the acoustic impedance between the tissue of interest and the surrounding tissue. The higher the contrast between the object and the background, the higher the variance of the pixel values of simulated images, thus the smaller the SNR of these images.

- **Size of the Object:** If the object of interest is small, the SNR of the resultant simulated image will be determined by the background. If the object is dominant, it will have more effect on the SNR of the simulated image. If the area of the object and the background region are comparable, the SNR of the image is further subjected to their contrast.

The last two factors can be combined as one factor, that is, the ratio of the total backscattered energy of the target tissue and the background tissue. With so many factors coming into play, it becomes almost infeasible to present simulated B-scan images which exhaust all possible scenarios with different combinations of the factors. Therefore in the following text we would rather concentrate on a few cases, particularly those of relevance to the real ultrasound images of the prostate.

Up until now, apart from the Rayleigh distribution, we have considered the K and Rician distribution. The K distribution has the limitation that it is applicable only if the SNR of the image is less than 1.9131. Similarly, the Rician distribution can only handle images with SNR greater than 1.9131. Therefore in the following text we also consider the following “all-around” distributions: Nakagami, Weibull, generalised Gamma and generalised Fisher-Tippett distribution which are applicable to images with any SNR values.

**Case 1: A Square Component**

Figure 4.13(a) shows one of 500 simulated B-scan images of a phantom with a specular square component at the centre. The relevant simulation setup is as follows:

- Scatterer density $d_s = 1.0$.
- No logarithmic compression applied.
- Backscattered echo amplitude\(^7\) of background tissue 0.5.
- Object/background contrast 5dB.
- The object (the square specular component) occupies 50% of the total area of the phantom, that is, the area of the object and the background are equal.

Since the SNR of these images are less than 1.9131, the Rician distribution is not applicable. A tentative experiment with 100 simulated images revealed that the Rayleigh,\(^7\)Normalised to [0,1] with 1 representing the largest amplitude.

72
Figure 4.13: Simulated B-scan image of a phantom with a square specular component. (a) Simulated image. (b) After histogram equalisation. (c) Histogram of the image in (a), together with the fitted K and generalised Gamma distribution.
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

<table>
<thead>
<tr>
<th>HYPOTHESISED DISTRIBUTION</th>
<th>$\chi^2$ TEST</th>
<th>K-S TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLE</td>
<td>CDF</td>
</tr>
<tr>
<td>K</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Generalised Gamma</td>
<td>8</td>
<td>130</td>
</tr>
</tbody>
</table>

Table 4.4: Number of rejections of the null hypothesis that the simulated B-scan images in Case 1 follow the K distribution or the generalised Gamma distribution.

Nakagami, Weibull and generalised Fisher-Tippett failed the goodness-of-fit tests. Thus the K and generalised Gamma distribution were fitted to each image by the methods of MLE, CDF and PDF, followed by chi-squared goodness-of-fit test at the $\alpha = 0.05$ level of significance and a K-S test at $\alpha = 0.01$.

Figure 4.13(b) shows the image after histogram equalisation. Figure 4.13(c) illustrates the histogram of the image shown in Figure 4.13(a), together with the fitted K and the generalised Gamma distribution. As we can see, these two models fit the histogram very well.

The first row of Table 4.4 lists the numbers of rejections of the null hypotheses that these images follow the K distribution. The numbers of rejections are fairly small for all cases except that the number of the method of PDF is higher than expected, but still acceptable for an experiment of 500 tests. The numbers of rejections of the null hypotheses that these images follow the generalised Gamma distribution are listed in the second row of the Table. As we can see, the number of rejections is higher than expected for the method of CDF and PDF in the case of chi-squared test, and the method of MLE in the case of K-S test. The numbers of rejections for the rest of them are reasonable. These test results provide strong support that the null hypotheses should be rejected, although the K distribution provides a better fit compared to the generalised Gamma distribution, as shown in Table 4.4 and Figure 4.14.

Figure 4.14(a) and (b) respectively compare the boxplots of the statistics of the chi-squared goodness-of-fit tests and the K-S tests of the two distributions. As we can see, the method of MLE is the best among the three methods in the sense of chi-squared goodness-of-fit, while the method of CDF is the best in the sense of K-S goodness-of-fit. This agrees with our previous observations.

Case 2: A Square Component (Logarithmic Compression Applied)

The simulation setup for this case is the same as that of the previous case except that the *ad hoc* logarithmic compression is applied. 500 images were simulated and the SNR of these images are all over 1.9131. These images generally look like the one shown in Figure 4.13(a) although marginally brighter. Therefore they are not shown here. A tentative experiment with 100 simulated images showed that the Rician, Nakagami, Weibull and generalised Gamma distribution are acceptable in terms of the goodness-of-fit tests. These four distributions were fitted to each image by the methods of MLE, CDF
Figure 4.14: Boxplots of the statistics of the goodness-of-fit tests of the fitted K and generalised Gamma (GA) distributions. (a) Chi-square tests. (b) K-S tests.
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

Figure 4.15: Histogram of one of the simulated images of a phantom which contains a square specular component, with *ad hoc* logarithmic compression applied. Together with the histogram are shown the fitted Rician, Nakagami, Weibull and generalised Gamma distribution. The curves of the latter three distributions are superimposed on one and another and almost indistinguishable.

<table>
<thead>
<tr>
<th>HYPOTHESISED DISTRIBUTION</th>
<th>$\chi^2$ TEST MLE</th>
<th>CDF</th>
<th>PDF</th>
<th>K-S TEST MLE</th>
<th>CDF</th>
<th>PDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rician</td>
<td>10</td>
<td>11</td>
<td>81</td>
<td>278</td>
<td>229</td>
<td>415</td>
</tr>
<tr>
<td>Nakagami</td>
<td>4</td>
<td>4</td>
<td>55</td>
<td>231</td>
<td>29</td>
<td>337</td>
</tr>
<tr>
<td>Weibull</td>
<td>4</td>
<td>6</td>
<td>78</td>
<td>256</td>
<td>63</td>
<td>383</td>
</tr>
<tr>
<td>Generalised Gamma</td>
<td>1</td>
<td>6</td>
<td>9</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 4.5: Number of rejections of the null hypothesis that the simulated B-scan images in Case 2 follow the specified distributions.
and PDF followed by chi-squared goodness-of-fit test at the $\alpha = 0.05$ level of significance and a K-S test at $\alpha = 0.01$.

Figure 4.15 shows histogram of one of these simulated images, together with the fitted Rician, Nakagami, Weibull and generalised Gamma distribution (using the method of MLE). As we can see, the latter three distributions all agree very well with the histogram. Besides, their curves are superimposed on one and another and almost indistinguishable, indicating that these three methods may be comparable. The fitted Rician distribution (dotted line) slightly deviates from the other three distributions.

Table 4.5 summarises the numbers of rejections of the null hypotheses that these images follow the specified distributions. As we can see, the generalised Gamma distribution has the least number of rejections in both the chi-squared test and K-S test. Thus it is safe to conclude that the generalised Gamma distribution is the most suitable distribution among the four models for describing the statistical properties of these images. The Nakagami distribution is also acceptable, since there are only 4 rejections with the method of MLE in the chi-squared tests, and 29 rejections with the method of CDF in the K-S tests. The Rician distribution yields the “worst” fit among the four distributions.

The boxplots of the chi-squared goodness-of-fit test statistics of the four distributions are shown in Figure 4.16(a). As we can see, the method of MLE is the best among the three parameter estimation methods in the sense of chi-squared goodness-of-fit. Figure 4.16(b) compares that of the K-S tests. Apparently the method of CDF is the best in the sense of K-S goodness-of-fit, which agrees with our previous observations.

**Case 3: A Small Square Component**

The setup in this case is the same as that of the previous case except that the object (the square specular component) occupies only 20% of the total area of the phantom (that is, the area ratio of the object and the background is 1:4). 500 images were simulated and the SNR of these images were all over 1.9131. One of these images is shown in Figure 4.17(a). Figure 4.17(b) is the image after histogram equalisation.

The Rician, Nakagami, Weibull, and generalised Gamma distribution were fitted to each image by the methods of MLE, CDF and PDF, followed by chi-squared goodness-of-fit test at the $\alpha = 0.05$ level of significance and a K-S test at $\alpha = 0.01$.

Figure 4.17(c) illustrates the histogram of the image shown in Figure 4.17(a), together
4.2. SIMULATED IMAGES AND STATISTICAL ANALYSIS

Figure 4.16: Boxplots of the statistics of the goodness-of-fit tests of the fitted Rician (RI), Nakagami (NA), Weibull (WE) and generalised Gamma (GA) distributions. (a) Chi-square tests. (b) K-S tests.
Figure 4.17: Simulated B-scan image of a phantom with a small square specular component. (a) Simulated image. (b) After histogram equalisation. (c) Histogram of the image in (a), together with the fitted Rician, Nakagami, Weibull and generalised Gamma distribution.
4.2 SIMULATED IMAGES AND STATISTICAL ANALYSIS

<table>
<thead>
<tr>
<th>HYPOTHESESED DISTRIBUTION</th>
<th>$\chi^2$ TEST</th>
<th>K-S TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLE</td>
<td>CDF</td>
</tr>
<tr>
<td>Rician</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Nakagami</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Weibull</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Generalised Gamma</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4.7: Number of rejections of the null hypothesis that the simulated B-scan images in Case 4 follow the specified distributions.

with the fitted Rician, Nakagami, Weibull and generalised Gamma distribution. Although visually these distributions yield a good fit, the goodness-of-fit tests revealed that the null hypotheses were rejected.

The numbers of rejections of the null hypotheses that these images follow the specified distributions are listed in Table 4.6. As we can see, except that the generalised Gamma distribution satisfies the K-S test when the method of CDF is used, all the other test results suggest that the null hypotheses should be rejected. Thus for this case, there is no suitable distribution to our knowledge for modelling the statistical properties of these images.

Case 4: A Small Hypoechoic Square Component

The setup in this case is the same as that of the previous case except that the object is hypoechoic. The object/background contrast is -5dB. 500 images were simulated and the SNR of these images were all over 1.9131.

Table 4.7 lists the numbers of rejections of the null hypotheses that these images follow the Rician, Nakagami, Weibull or the generalised Gamma distributions. As we can see, the Rician distribution is appropriate for modelling the statistical properties of these images, although the Nakagami, Weibull and generalised Gamma distributions are also reasonable. Compared to the previous case, as we have seen, a “flip” in the object/background contrast have significantly changed the statistical properties of the images.

Case 5: Three Circular Components

In this case there are three circular specular components. The contrast between the three circular objects and the background are respectively -5dB, 6.5dB and 6.5dB. Note that this setup has a special implication to the real ultrasound images, that is, the two small bright discs (6.5dB) can represent some specular medium with strong reflections, while the big dim disc (-5dB) can then represent the shadow created by the specular medium.

One of 500 simulated images of this phantom is shown in Figure 4.18(a), and its histogram-equalised version in Figure 4.18(b). Its histogram, together with the fitted Nakagami, Weibull and the generalised Gamma distributions, is shown in Figure 4.18(c). Although it appears that these distributions fit reasonably well to the histogram, the
Table 4.8: Number of rejections of the null hypothesis that the simulated B-scan images in Case 5 follow the specified distributions.

<table>
<thead>
<tr>
<th>HYPOTHESISED DISTRIBUTION</th>
<th>(\chi^2) TEST</th>
<th>K-S TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MLE</td>
<td>CDF</td>
</tr>
<tr>
<td>Nakagami</td>
<td>493</td>
<td>496</td>
</tr>
<tr>
<td>Weibull</td>
<td>495</td>
<td>497</td>
</tr>
<tr>
<td>Generalised Gamma</td>
<td>461</td>
<td>489</td>
</tr>
</tbody>
</table>

goodness-of-fit tests showed that the null hypotheses were rejected.

The mean and median of the SNRs of the 500 images are respectively 1.9033 and 1.9035, and the standard deviation is 0.0138. Since the SNRs are not always above 1.9131, only the three “all-around” distributions, the Nakagami, Weibull and the generalised Gamma distributions, were fitted to the histogram of the images.

The results of the subsequent goodness-of-fit tests are summarised in Table 4.8. As we can see, the Nakagami distribution satisfies the K-S test when the method of CDF or PDF is used. The Weibull and the generalised Gamma distribution satisfy the K-S test when the method of CDF is used. All the other test results suggest that the null hypotheses should be rejected. Thus for this case, there is no “ideal” fit distribution for modelling the statistical properties of these images. This suggests that the presence of the hyperechoic and hypoechoic (or anechoic) medium can significantly change the statistical properties of B-scan ultrasound images, making it almost a “mission impossible” to find an “all-around” distribution that fits perfectly to the histogram of the images.

4.3 Statistical Analysis of Real Ultrasound Images

The computer simulation of B-scan ultrasound imaging and the subsequent analysis of the statistical properties of the simulated images in the previous section inspired us to find a distribution model that best fits the ultrasound images of the prostate in this section.

Figure 4.19(a) shows the histogram of image “us1”. As we can see, there are two artifacts with this histogram, which have to be addressed before various distributions can be fitted to the histogram.

- The first problem is that there are many “null” bins (the pixel counts of which are all zero) alternating the non-zeros bins. This is due to the fact that the Proscan® Plus imaging system has only 128 gray scale levels. The ultrasound images produced by this system were digitised in 256 levels of gray scale, thus creating a “stretched” histogram. These “null” bins have to be removed before fitting the various distributions to the histogram. (Interpolation of the histogram is not appropriate since this would create some otherwise nonexistent “new” pixels, which could not be justified.

By “ideal” we mean the fitted distribution satisfies both the chi-squared and the K-S goodness-of-fit tests.
4.3. STATISTICAL ANALYSIS OF REAL ULTRASOUND IMAGES

Figure 4.18: Simulated B-scan image with three specular coherent components. (a) Simulated image. (b) After histogram equalisation. (c) Histogram of the image in (a), together with the fitted Nakagami, Weibull and generalised Gamma distribution.
Figure 4.19: Histogram of image “us1”. (a) The original histogram. Note the “null” bins. (b) “Null” bins removed. (c) Signal-free regions removed and normalised to relative frequencies.

83
4.3. **STATISTICAL ANALYSIS OF REAL ULTRASOUND IMAGES**

![Histogram of image “us1” and the fitted distributions.](image)

**Figure 4.20**: The histogram of image “us1” and the fitted Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions.

“Null” bins are removed by searching the histogram for bins whose pixel count is nonzero and then keeping only the nonzero bins.

- Another problem is the peak at the left (low gray scale levels) of the histogram. An inspection of all the ultrasound images revealed that the peak is due to the signal-free regions (the dark regions at the four corners due to geometry transform) of the images. The peak is “consistent” across all the images in that a threshold can be set to remove it from the histogram. This is equivalent to defining a region-of-interest or mask which excludes the signal-free area. The threshold is set to 0.18 (normalised gray scale) for all images.

The histogram after the “null” bins are removed is shown in Figure 4.19(b). The image is further thresholded at 0.18 to remove the signal-free regions. The histogram is then normalised to relative frequencies. The final form of the histogram is shown in Figure 4.19(c).

After the artifacts of the histogram are removed, various distributions can be fitted to the histogram. Since the SNR of the image is over 1.9131, the K distribution is not applicable. Thus the Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions were fitted to the image by the methods of MLE, CDF and PDF, followed by chi-squared goodness-of-fit test at the $\alpha = 0.05$ level of significance and
a K-S test at $\alpha = 0.01$.

Figure 4.20 shows the results of fitting the distributions to the histogram of image “us1”. As we can see, the fitted Rician, Nakagami, Weibull and generalised Gamma distributions only agree reasonably well with the right half of the histogram, but fail to capture the deep ascent at the left half of the histogram. The generalised Fisher-Tippett distribution picks up the left deep ascent very well, but not the case for the right half. Figure 4.21 and Figure 4.22 respectively illustrates the results of fitting the distributions to the histogram of image “us6a” and “us6b”. As we can see, the generalised Fisher-Tippett distribution provides the best fit among the five distributions. However, all these distributions failed the goodness-of-fit tests.

Figure 4.23(a) illustrates the boxplots of the chi-squared goodness-of-fit test statistics of the Rician (RI), Nakagami (NA), Weibull (WE), generalised Gamma (GA), and generalised Fisher-Tippett (FT) distributions (the method of MLE was used for parameter estimation). Figure 4.23(b) shows the boxplots of the K-S test statistics of the five distributions (the method of CDF was used for parameter estimation). Both figures suggest that the generalised Fisher-Tippett distribution provides the best fit to the histograms of the images. However, a perfect fit could not be found.

The reason why the generalised Fisher-Tippett distribution failed the goodness-of-fit tests can be attributed to the “spikes” along the histograms of the images. As we can see in Figure 4.20-4.22, the rightmost spikes are due to the bright calibration markers in the original images. Although this spike (ultimately the pixels corresponding to the markers) can be removed by thresholding, there are other spikes which are not easy to isolate. These spikes comprise the dominating factor that caused the failure of the generalised Fisher-Tippett distribution in terms of the goodness-of-fit. There is no way to circumvent these spikes, no matter what distribution model is fitted to the histogram.

### 4.4 Conclusion

In this chapter we presented a hypothetical B-mode ultrasound imaging system and various simulated images for different scenarios. The speckle amplitudes of a simulated B-scan image of a homogeneous phantom normally follow the Rayleigh distribution, but skew away to become pre-Rayleigh (the K-distribution) when the number of scatterer per resolution cell is sparse, or become post-Rayleigh (the Rician distribution) when the ad hoc logarithmic compression is applied.

The K distribution can only handle images whose SNRs are less than 1.9131, while the Rician distribution has the limitation that it is applicable only if the SNR of the image is greater than 1.9131. Therefore more distribution models were introduced, including the Nakagami, Weibull, generalised Gamma which are applicable to images with any SNR.

Generally a distribution estimated by the method of MLE provides the best fit in the sense of chi-squared goodness-of-fit, whereas the method of CDF yields the least K-S test.
4.4. CONCLUSION

Figure 4.21: The histogram of image “us6a” and the fitted Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions.

Figure 4.22: The histogram of image “us6b” and the fitted Rician, Nakagami, Weibull, generalised Gamma, and generalised Fisher-Tippett distributions.
Figure 4.23: Boxplots of the statistics of the goodness-of-fit tests of the fitted Rician (RI), Nakagami (NA), Weibull (WE), generalised Gamma (GA) and generalised Fisher-Tippett (FT) distributions. (a) Chi-square tests (MLE). (b) K-S tests (CDF).
statistics since it is cdf “oriented”. In other words, a distribution estimated by any one of the three methods cannot address the two test criteria equally well at the same time. (A perfect fit always satisfies the chi-squared and K-S goodness-of-fit tests.)

When specular components are present in a phantom, the statistical properties of simulated images are not as predictable as that of a homogeneous phantom anymore since there are many factors that affect the statistical properties of simulated images. As we have seen, the speckle pattern is determined not only by the imaging system but also by the underlying structure of the tissue being scanned. It is not uncommon that an ideally fit distribution model that satisfies both the chi-squared and the K-S goodness-of-fit tests cannot be found.

Statistical analysis of the simulated images inspired the effort of finding the underlying distribution that models statistical properties of the ultrasound images of the prostate. We empirically proposed the generalised Fisher-Tippett distribution and found it promising in modelling the ultrasound images, but it only provides a best fit among the distribution models in comparison, not a perfect fit in terms of goodness-of-fit.

Nevertheless, computer simulation of B-scan ultrasound images provides a viable test-bed for validation of our speckle reduction scheme later in Chapter 6. Before we present the speckle reduction scheme, let us review the current literature on speckle reduction in Chapter 5.
Chapter 5

Speckle Noise Reduction: A Survey

Since speckle is an undesired artifact that plagues interpretation of optical, synthetic aperture radar (SAR), and ultrasound images, extensive efforts have been devoted to counteracting this adversity. These attempts generally fall into two categories: those at the system level and those at the image level.

By “system level” we mean pre-image-formation, that is, an attempt is made to suppress speckle before final images are formed by the imaging system or before the images are presented to the display device for visual inspection by users. We refer to the off-line or post-image-formation image processing techniques as “image level” speckle reduction.

We first survey in section 5.1 system level speckle reduction techniques such as compound scanning and Seggie’s phase-based deterministic approach, and then focus on off-line image processing techniques in the remaining text. These techniques include classical adaptive filters, wavelet methods and geometry driven diffusion and other methods.

5.1 System Level Speckle Noise Reduction

5.1.1 Compound Scanning

When an object is scanned under different conditions, the speckle patterns are quite different. However, if the same object is insonified repeatedly under relatively the same conditions (say same transducer aperture and same transmitted pulse width, but the transducer is moved to a slightly different position), statistically identical speckle patterns of the resultant images are expected, despite their random appearance. In other words, although the pixel values of the same point across the two images are usually quite different, their statistical properties as a whole are the same. This is the underlying motivation of compound scanning.

In compound scanning, the object is scanned from different positions and/or different angles of view (spatial compounding [21, 73]) or at different pulse frequencies (frequency
compounding [74–76]). Since the backscattered wavelets within the resolution cells have different spatial relationships (relative to the transducer), the echo signal from the same volume element of the object will fluctuate randomly from image to image. Therefore, either taking the average value (compound scan with averaging [77, 78]) or the maximum of the echo amplitude (compound scan with maximum amplitude writing [21]) of the volume elements will destroy the coherence of the echo signal and reduce the speckle.

Speckle reduction via spatial compounding with averaging is equivalent to applying a low-pass moving average filter to the images, since in spatial compounding speckle across different images obtained for the same target can be considered uncorrelated, and averaging these uncorrelated images resembles taking the average of the neighbouring pixels of a single image. It has been shown that the improvement in SNR of compound scanning with maximum amplitude writing is comparable to that of compound scanning with averaging [21]. Compound scanning with averaging is equivalent to multi-look incoherent averaging techniques commonly used for restoration of SAR imagery.

Galloway [79] proposed a frequency diversity process for speckle reduction, in which the received echo spectrum was separated into sub-spectra. Upon averaging the resultant image lines, the speckle was decorrelated and reduced. This method essentially could be regarded as a variation of frequency compounding.

Recently a new compounding technique referred to as strain compounding was proposed by Li and Chen [80], which reduces speckle variations by exploiting the decorrelation between images under different strain states. It was found that externally applied forces result in tissue motion and different strain states. The in-plane (axial or lateral) motion can be estimated and corrected so that images generated under different strain states are spatially matched. The out-of-plane motion is not corrected, therefore the images are partially correlated. Reduction of speckle is thus achieved by incoherently averaging these partially correlated images. In essence this technique falls in the class of spatial compounding.

### 5.1.2 Seggie’s Phased-Based Deterministic Approach

Seggie [70, 81] proposed a rather different approach towards the problem of speckle reduction, which is based on the analysis of the instantaneous frequency (time derivative of the signal phase) of the echo signal. A unique feature of this approach is that it is aimed at real-time implementation and can be applied to ultrasound radio frequency (RF) signal.

The instantaneous frequency of a signal is defined by

$$\phi' = \frac{d\phi}{dt}$$  \hspace{1cm} (5.1)

where

$$\phi = \arctan \frac{\hat{A}}{A},$$  \hspace{1cm} (5.2)

and $\hat{A}$ and $A$ are the constituent parts of the analytical signal defined in Equation (4.8).
SPECKLE NOISE REDUCTION: A SURVEY

(here we omit the independent variable \( x,y \) to emphasise the temporal nature of \( A \) and indicate that it is one of the A-mode echo signals rather than an ensemble of them).

It was found that the instantaneous frequency \( \phi' \) exhibits very large excursions at locations where echoes from the scatterers overlap, which results in destructive interference. Therefore, speckle can be located and marked by identifying these large \( \phi' \) fluctuations. To reduce the speckle, a provisional simple filling-in method was proposed, in which the A-mode envelope values in a small neighbourhood around the speckle marker were replaced by the average of the two adjacent envelope maxima on either side of the speckle marker. It was shown that this speckle reduction technique, even with the filling-in method in its provisional form, greatly improved image quality and yielded better results than other methods when assessed on some quantitative indices [70, 81]. Other speckle reduction methods based on the instantaneous frequency can be found in [82, 83].

5.1.3 Hokland’s Harmonic Imaging

Hokland and Taxt [71, 84] proposed a speckle reduction method called harmonic imaging (HI) based on harmonic oscillator models. The authors modelled the backscattered ultrasound quadrature signal as the sum of three complex components, namely,

- the non-resolvable component corresponding to all randomly positioned diffuse scatterers;

- the sub-resolvable component, which is attributed to quasi-periodic structures, the spatial variations of which become resolvable relative to the ultrasound pulse length;

- the resolvable component, which results from specular surfaces and quasi-periodic structures that are resolvable relative to the ultrasound pulse.

They deemed that speckle is mostly due to the sub-resolvable component, which gives rise to most of the destructive interference, and can be removed simply by differentiating the corresponding backscattered demodulated quadrature signal. The sum of the signal due to the remaining two components is then modelled by two orthogonal and independent harmonic oscillators. The gray scale levels of the (denoised) images are then determined by estimating the total energy of the two oscillators.

The HI method can be regarded as a nonlinear split spectrum algorithm [71]. However, the split of spectrum is accomplished in the spatial domain through incorporation of prior information about the composition of the backscattered signal, instead of in the frequency domain like that of the ordinary split spectrum algorithms such as [79]. The authors believed that the HI image restoration algorithm compares favourably with the ordinary split spectrum algorithms.
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

5.1.4 Comments on System Level Speckle Reduction

The speckle reduction methods at the system level are not without their own limitations. For compound scanning, the transducer has to be moved by about half its width between measurements. Another problem is that it has to be assured that the sound beam insonifies exactly the same object every time, and it requires an experienced operator to satisfy this condition (and then only approximately) [21].

Moreover, speckle reduction in compound scanning is not achieved without a cost. An undesirable side effect of averaging or maximum amplitude writing is the reduction in resolution and loss of small texture patterns. Recently Evans and Nixon [85,86] proposed a two-dimensional least mean square filter, which was configured as a motion-adaptive temporal filter for a time sequence of ultrasound images, and aimed at eliminating the blurring associated with direct averaging. However, a disadvantage of this method is its computational cost.

For Seggie’s deterministic approach [70,81] and Hokland’s HI method [71,84], one has to have access to the radio-frequency (RF) A-mode scan lines of the imaging system. Once the envelope detection is applied, the phase information carried by the A-mode lines is lost and not recoverable anymore. Unfortunately, it is usually not possible for an end-user (consumer) to have access to the internals of a commercial ultrasound imaging system. Therefore, in the following text, we will focus on speckle noise reduction methods at the image level, that is, “off-line” single-frame image processing techniques, although we deem that “on-line” speckle reduction at the system level is quite promising, desirable and warrants further research because of the real-time nature of clinical ultrasound, as supported by the recent developments reported in [80].

5.2 Image Level Speckle Noise Reduction

Since speckle is the prime culprit that masks off meaningful texture information and degrades contrast and resolution of optical, SAR and ultrasound images, many speckle filtering techniques have been proposed in these areas, especially in the SAR imaging community. These methods can be classified as

- Classical adaptive filtering methods, which are based on local statistics of the images, including the Lee filter [87–89], Kuan filter [90], Frost filter [91], and Loupas’ adaptive weighted median filter (AWMF) [92];

- Wavelet methods [93–98], especially Donoho’s seminal work on soft thresholding (shrinkage) [99,100] of coefficients of the wavelet transform;

- Geometry driven diffusion, which is based on nonlinear partial differential equations for smoothing images on a continuous domain [72,101].
5.2.1 Classical Adaptive Filtering Methods

Common to the adaptive noise filtering techniques described below is the use of the local statistics of the images. Underpinning these methods are the various noise models assumed, including additive [87, 89, 102], signal independent multiplicative [87–89, 91, 102–107], and signal dependent multiplicative noise models [92, 95, 107, 108].

Noise Models

A commonly encountered additive noise model [87, 89, 102] is described by

\[ x = s + n \]  \hspace{1cm} (5.3)

where \( x \) is the observed image, \( s \) is the noise-free image and \( n \) is the additive noise, usually modelled as a normal random variable with zero mean and standard deviation \( \sigma_n \), that is \( n \sim N(0, \sigma_n) \). It was shown that the speckle noise can be approximately regarded as Gaussian additive noise when the SAR imaging system has a finite aperture and the images are logarithmically transformed [93, 109].

It is well known to the SAR imaging community that speckle has the properties that resemble that of an uncorrelated multiplicative noise in the sense that the noise level is proportional to the local average gray levels. In other words, the speckle is more pronounced in a “bright” area than in an area of low average intensity. The multiplicative noise model is described by [87–89, 91, 102–107, 110]

\[ x = sn \]  \hspace{1cm} (5.4)

where the noise-free image \( s \) and noise \( n \) are assumed to be two independent random processes, and it is commonly accepted that the speckle follows the Rayleigh distribution (see Equation (3.2)) when appropriate conditions hold (see Section 3.1). The Rayleigh speckle is multiplicative in the sense that its standard deviation is proportional to its mean (see subsection 3.1.1).

Apart from the above two noise models, some researchers have elected to model speckle noise as signal dependent and multiplicative [92, 95, 107, 108], that is,

\[ x = s + \sqrt{s} n. \]  \hspace{1cm} (5.5)

Other speckle scene models for SAR imagery include the product model and the multiply stochastic model [111]. Touzi [111] studied the multiplicative and product speckle models and reviewed speckle filtering techniques in the context of estimation theory. He demonstrated that speckle filtering can only be effective in locally stationary scenes.

Based on the above noise models various adaptive speckle filtering techniques have been developed. A comparison of some of these classical techniques can be found in [112]. A comparative study between a complex wavelet coefficient shrinkage filter and some
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

classical speckle filters is presented in [113]. Hervet [114] and his colleagues also did a comparison between wavelet methods and classical adaptive statistical speckle filtering methods.

Lee’s Filter

Initially assuming an additive noise model and treating speckle suppression as a problem of minimum mean-square error (MMSE) estimation of the noise-free signal \( s \) from the observed signal \( x \), Lee [87, 88] proposed a local statistics based adaptive filter of the form

\[
\hat{s} = \bar{x} + k(x - \bar{x}) \tag{5.6}
\]

where \( \hat{s} \) is the MMSE estimator of the noise-free signal \( s \), that is, the denoised signal. \( \bar{x} \) is the mean value of the pixels within a \( N \times N \) filter window \( W \) (or referred to as the kernel in some literature). \( x \) represents the current observed pixel at the center of the filtering window. \( k \) is an adaptive weighting coefficient determined by

\[
k = 1 - \frac{\sigma_n^2}{\sigma_x^2} \tag{5.7}
\]

where \( \sigma_n^2 \) is the variance of speckle and assumed to be known a priori. In practice it can be estimated from a region of fully formed speckle in the image. \( \sigma_x^2 \) is the variance of the pixels within the filter window

\[
\sigma_x^2 = \frac{1}{N^2 - 1} \sum_{x \in W} (x - \bar{x})^2. \tag{5.8}
\]

Intuitively, in a “flat” area dominated by speckle, \( k \) tends to 0 and \( \hat{s} \) takes the mean value \( \bar{x} \). This is equivalent to applying a uniform smoothing filter, therefore speckle is suppressed to the greatest possible extent. In contrast, in an area abundant in features such as edges and lines, \( \sigma_x^2 \gg \sigma_n^2 \), therefore \( k \rightarrow 1 \) and \( \hat{s} \rightarrow x \), that is, smoothing is kept to a minimum and the pixel tends to retain its original value. Therefore features like edges are preserved to the greatest possible extent.

For multiplicative noise, Lee [87, 89] derived the optimal linear approximation of Equation (5.4) by virtue of the first order Taylor series expansion about \((\bar{s}, \bar{n})\)

\[
x = \bar{n} s + \bar{s}(n - \bar{n}). \tag{5.9}
\]

After applying the MMSE criterion to this linear approximation and assuming unit mean for the noise \( (\bar{n} = 1) \), Lee obtained a filtering algorithm of the same form as Equation (5.6) but with a different adaptive weighting coefficient which changes to

\[
k = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} \tag{5.10}
\]
where $\bar{s}$ and $\sigma^2_s$ are respectively the a priori mean and variance of the noise-free signal, and can be approximated by

$$\bar{s} = \frac{\bar{x}}{n} = \bar{x}$$

and

$$\sigma^2_s = \frac{\sigma^2_x + \bar{x}^2}{\sigma^2_n + n^2} - \bar{s}^2 = \frac{\sigma^2_x - \bar{x}^2 \sigma^2_n}{\sigma^2_n + 1}$$

(5.12)

where $\sigma_n$ can be estimated by

$$\sigma_n = \frac{\sigma_x}{\bar{x}}$$

(5.13)

from a feature-free area with fully developed speckle. The author also extended the filtering algorithm to deal with images corrupted by both additive and multiplicative noise [87].

The effectiveness of Lee’s filter is greatly affected by the filter window size. Tradeoff has to be made between noise reduction and preservation of subtle details of the images in selecting the window size, which is usually determined by the user on a trial and error basis. The author found that a $7 \times 7$ window is a good choice for the SAR images used in his experiments [87–89].

Kuan’s Filter

Kuan [90] proposed a filtering method similar to Lee’s in a sense that image restoration or speckle reduction is approached as a MMSE estimation problem. Based on a non-stationary mean and non-stationary variance image model, Kuan imposed a linear constraint on the estimator structure and derived a local linear MMSE filter, which is of the same form as Lee’s filter for additive noise.

For multiplicative noise, the noise model in Equation (5.4) was transformed into a signal plus signal-dependent additive noise form and the variance of the noise term was derived. The linear constraint was applied to the estimator structure rather than the nonlinear observation model. An optimal linear filter of the same form as Equation (5.6) was then derived as a generalisation of Lee’s filter for multiplicative noise. Compared to that of Lee’s filter (see Equation (5.10)), the adaptive weighting coefficient of Kuan’s filter has an additional term in the denominator

$$k = \frac{\sigma^2_s}{\sigma^2_s + \sigma^2_n(\bar{x}^2 + \sigma^2_s)}$$

(5.14)

As we can see from the above, Kuan’s filter and Lee’s filter are similar in a sense that the structure of the adaptive speckle filter is separated from the estimation of local statistics of the images and remains unchanged for additive and multiplicative noise. The advantage of keeping filter structure “fixed” is that various methods can be used or extended to estimate local statistics for different scenarios. The only difference between these two methods lies just in their different approaches to local statistics estimation.
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

Frost’s Filter

Speckle reduction was also treated as a problem of MMSE estimation by Frost [91], on the basis of the multiplicative noise model and local statistics of the images, like the Lee and Kuan filters. Under the assumptions of stationary image data, an adaptive filter was derived, which provides the MMSE estimate in homogeneous areas. A decay coefficient $\alpha$ which varies with position was further introduced to address the non-stationary aspects of real SAR images.

Frost’s filter can be described by

$$\tilde{x}_c = \sum_{x \in W} K_0 \alpha e^{-\alpha d} x$$

(5.15)

where $\tilde{x}_c$ is the current filtered pixel, $K_0$ is a normalising constant, and $d$ is the distance between $x$, the pixels in the filter window $W$ and $x_c$, the current pixel being filtered at the centre of $W$. The adaptive decay coefficient $\alpha$ is estimated from the local statistics in the filter window $W$ as

$$\alpha = K_1 \sqrt{\frac{\sigma^2}{\bar{x}^2}}, \quad x \in W$$

(5.16)

where $K_1$ is a constant related to $\sigma^2_n / \bar{n}^2$, the ratio of the variance to the square of the mean of speckle noise.

Intuitively, in a “flat” area without features like edges or lines, $\alpha$ tends to be small, and the filter behaves more like a moving average filter. On the contrary, prominent edges result in a large $\alpha$ and the filter has a tendency to preserve them. Therefore the Frost filter can be regarded as an adaptive weighted mean filter.

Loupas’ Adaptive Weighted Median Filter

As we have seen, the Lee, Kuan, and Frost filters are all based on local statistics such as the mean and variance of the images. Another possibility in suppressing speckle is to make use of local order statistics such as the median. These filters are nonlinear because of the ordering process involved, and are known as order statistics filters (OSF) [92, 115–117], among which the median filter [117, 118] is by far the most popular due to its simplicity in implementation. It has been shown that the median filter is extremely effective in suppressing impulsive noise such as “salt-and-pepper” noise [118]. Since speckle is not of an impulsive nature (for example, the gray scale values of Rayleigh speckle span the whole “spectrum” of the histogram, see Figure 3.1), the median filter is not very effective in suppressing speckle noise.

Based on a signal dependent and multiplicative noise model, Loupas [92] proposed an adaptive weighted median filter (AWMF), which is an extension to Brownrigg’s weighted median filter [119]. The weighted median filter operates as follows. For a data array of $N$ elements, $\{x_i, \ 1 \leq i \leq N\}$, the extended array $\{x_i^E\}$ is constructed by repeating each

96
element $x_i k_i$ times,

$$\{x^E_i\} = \{x_1, \ldots, x_1, x_2, \ldots, x_2, \ldots, x_N, \ldots, x_N\}$$ (5.17)

where $k_i$ is the corresponding weight for $x_i$. The filter output is then obtained as the median of the extended data array $\{x^E_i\}$

$$\tilde{x} = \text{Med}\{x^E_i\}$$ (5.18)

where $\tilde{x}$ denotes the filter output and $\text{Med}\{\cdot\}$ represents taking the median. By varying the weights $k_i$, we can balance between noise suppression and signal preservation. In particular, Loupas [92] designed a space-varying weight as

$$k_{i,j} = \left\lfloor k_c - c d_{i,j} \frac{\sigma^2_x}{\bar{x}} \right\rfloor, \quad -N \leq i, j \leq N$$ (5.19)

where $\lfloor x \rfloor$ denotes taking the nearest integer to $x$ if $x > 0$, or zero if $x < 0$. $i, j$ represent the spatial coordinates within a filter window of size $(2N+1) \times (2N+1)$. $k_c$ is the weight for the centre pixel. $c$ is a scaling constant. $d_{i,j}$ is the distance between the pixel at $(i, j)$ and the centre pixel. $\sigma^2_x$ and $\bar{x}$ respectively denotes the local variance and mean of the pixels within the filter window.

As we can see, the weight $k_{i,j}$ decreases as the pixel moves away from the centre of the filter window and the rate of decrease is determined by the local statistics within the filter window. Intuitively, for a region with features like a boundary, the ratio of the local variance to the local mean is large, thus emphasis is put on the centre pixel, and the filter tends to keep the original pixel value and fine image detail can be preserved. For feature-free areas where gray scale fluctuations are only attributed to speckle noise, the weights are nearly uniformly “spread” out and maximum noise reduction is achieved.

The extent to which noise is reduced is determined by the window size, and the ability of preserving image detail is jointly controlled by the scaling constant $c$ and central weight $k_c$. These parameters have to be selected experimentally in practice. Moreover, computational complexity is another disadvantage of this filter [92].

**Other Adaptive Filtering Methods**

Apart from the filters introduced above, there are other notable adaptive speckle filters reported in the literature, including

- the truncated median or mode filter first proposed by Davies [120] and applied to ultrasound images by Evans and Nixon [121, 122], which takes advantage of the ranking order of the mode, median and mean to estimate the mode for the underlying distribution of ultrasound images, but may yield erratic results around the edges;
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

- Lee’s sigma filter [123–125], which is basically an adaptive mean filter, but only includes the pixels whose gray scale values are within two standard deviations from the *a priori* mean for computation of the mean value;

- the enhanced Kuan and Frost filters proposed by Lopes, Touzi, and Nezry [105], who elected to divide an image into small regions of three classes based on the variation coefficients and treat them differently: for homogeneous areas maximum speckle reduction is achieved by applying a uniform mean filter, for heterogeneous areas the enhanced Kuan or enhanced Frost filter is applied, and for areas with isolated point targets, no filtering is applied;

- Karaman and his coworkers’ adaptive speckle suppression filter [108], whose kernel size and shape are not fixed, but change adaptively through a local statistics based region growing technique;

- MAP filter based on the maximum *a posteriori* (MAP) criterion, which was first proposed by Kuan [104], and later extended by Lopez and his colleagues [126, 127], who elected to employ the Gamma and symmetrical Beta distributions as more realistic *a priori* statistical models for the underlying scene intensity, rather than the Gaussian used by Kuan [104];

- an adaptive mean filter recently developed by Dong, Milne, and Forster [128], which essentially relies on a Laplacian of Gaussian operator to differentiate regions with edges from those without edges by identifying zero-crossings;

- an adaptive speckle reduction filter proposed by Dutt and Greenleaf [27, 40], which is based on the K distribution statistical model for the echo envelope of ultrasound images, and accounts for the logarithmic compression of the echo envelope commonly used in clinical ultrasound imaging system.

**Comments**

Traditional noise reduction methods like Wiener or Kalman filtering [117, 118] require a statistical model for the images. For example, techniques based on Wiener filtering require knowledge of the power spectrum of the signal, which in turn relies on evaluation of autocorrelation functions of the noise and original noise-free signal. The autocorrelation function of the noise is analytically available if speckle is assumed to be fully developed. However, that of the noise-free signal is usually unknown, although some iterative procedures were proposed as a compromise [129, 130]. Since it is virtually impossible to obtain a simple and accurate model for the speckled images, these methods are believed to be not very effective in removing the speckle noise [89]. Another disadvantage of these techniques lies in that the non-stationary nature of the images is not accounted for, and they are inherently not spatially adaptive [131].
All of the adaptive filtering techniques introduced above have a common basis, that is, the local statistics such as the mean and variance or local order statistics such as the median. Moreover, most of them share the same filter structure (Equation (5.6)), and the difference between them is in the determination of the adaptive weighting coefficient \( k \) from local statistics.

Every noise reduction algorithm must face two conflicting requirements: noise reduction and feature preservation. The adaptive filtering methods are very effective in suppressing the speckle in homogeneous areas. Compromise has to be made between speckle elimination and image detail preservation around areas with important image features. Most of these methods are not very effective around these areas.

The effectiveness of these methods generally depends on the filter window size, which has to be determined empirically. It must be sufficiently large for accurate estimation of noise from local statistics, and it also has to be small enough for preservation of resolvable structures. Karaman [108] addressed this problem by using windows of arbitrary shape and size depending on local image content, but a disadvantage of this method is the possibility of producing some high-frequency artifacts in the filtered images.

Another concern for most of these adaptive filtering methods is the selection of control parameters. Like the window size, these parameters have to be adjusted on a trial and error basis for different images. Computational complexity is also a big concern for some of these methods despite the fact that each pixel can be processed independently of other pixels, making these methods naturally suitable for parallel processing [89].

### 5.2.2 Wavelet Methods

Thanks to its virtues of space and scale localisation, wavelet analysis has been vigorously investigated in the last two decades and successfully applied to many problems including noise reduction. Wavelet denoising methods usually involve manipulation of the coefficients of the wavelet transform of a signal at fine scales, which are usually dominated by noise.

#### Wavelet Shrinkage

One of the most recognised noise reduction approaches using wavelets is the wavelet shrinkage method developed by Donoho and Johnstone [99, 100, 132]. The authors assumed an additive noise model, where the noise is presumably independent and identically distributed Gaussian noise with mean zero and known variance. This method is based on the reasonable assumption that only a few large wavelet coefficients really contain information about the unknown signal \( s \), compared to the large number of small coefficients which can be attributed to the noise. The wavelet shrinkage method for recovering signals from noisy data is summarised as follows:

First a wavelet transform is applied to the observed noisy data \( x \) to yield coarse scaling coefficients \( a \) and wavelet coefficients \( w \). Since the discrete wavelet transform is essentially
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

A linear operation, the additive noise model still holds in the wavelet domain \([98,133]\), that is,

\[
\mathbf{w} = \mathbf{w}_s + \mathbf{w}_n
\]  

(5.20)

where \(\mathbf{w}_s\) and \(\mathbf{w}_n\) respectively represents the wavelet coefficients due to the noise-free signal \(s\) and the additive Gaussian noise \(n\). Thus the problem of denoising is reduced to estimate \(\mathbf{w}_s\) from the noisy observation \(\mathbf{w}\). The scaling coefficients \(\mathbf{a}\) which represent the gross structure of the signal are left intact. To suppress the noise, Donoho and Johnstone \([99, 100, 132]\) proposed to truncate the wavelet coefficients \(\mathbf{w}\) in amplitude by either a soft thresholding (shrinkage) or a hard thresholding rule. Finally, an inverse wavelet transform on the scaling coefficients \(\mathbf{a}\) and the thresholded wavelet coefficients \(\hat{\mathbf{w}}_s\) yields an estimate of the signal \(\hat{s}\).

The soft thresholding shrinks (hence the term “wavelet shrinkage”, see Figure 5.1 on page 104) the wavelet coefficients \(\mathbf{w}\) toward 0 by an amount \(T\) (the threshold), that is,

\[
\hat{\mathbf{w}}_s = \begin{cases} 
\mathbf{w} - T & \mathbf{w} \geq T \\
0 & |\mathbf{w}| < T, \\
\mathbf{w} + T & \mathbf{w} \leq -T 
\end{cases}
\]  

(5.21)

and the hard thresholding only keeps the original values of the prominent wavelet coefficients \(\mathbf{w}\) whose amplitude are above \(T\) (see Figure 5.1 on page 104)

\[
\hat{\mathbf{w}}_s = \begin{cases} 
\mathbf{w} & |\mathbf{w}| \geq T \\
0 & |\mathbf{w}| < T.
\end{cases}
\]  

(5.22)

Generally hard thresholding creates discontinuity on the wavelet coefficients and thus yields less smooth fits, but preserves important features. On the other hand, soft thresholding provides smoothness, but at the cost of shrinkage of the wavelet coefficients of genuine features.

The choice of the threshold \(T\) is crucial to the success of denoising via wavelet shrinkage and various methods have been proposed to estimate the threshold. In their seminal work Donoho and Johnstone \([99]\) suggested a “universal” threshold

\[
T = \sigma_n \sqrt{2 \log(N)}
\]  

(5.23)

where \(N\) is the sample size (length) of the signal and \(\sigma_n\) is the standard deviation of the noise. This threshold is optimal for restoration of the original signal \(s\) in the minimax sense (that is, minimising the maximum error over all possible signal of length \(N\) and for a wide range of loss functions). In practice, however, the noise level is usually unknown, and Donoho \([100]\) proposed estimating \(\sigma_n\) from the wavelet coefficients of the finest scale.
by

\[ \hat{\sigma}_n = \frac{\text{MAD}(w_1)}{0.6745} \]
\[ = \frac{\text{Med}(|w_1 - \text{Med}(w_1)|)}{0.6745} \]

(5.24)

where \( \text{MAD}(w_1) \) is the median absolute deviation (from the median) of the wavelet coefficients of the finest scale \( w_1 \).

Although the “universal” threshold given in Equation (5.23) is simple and efficient in removing the noise, it takes virtually no account of the real signal, and thus only provides a probabilistic upper bound on the amplitude of the noise in the wavelet coefficients over all scales [133]. Therefore it should be applied with caution since it could over-smooth the signal. Usually it has to be fine-tuned on a scale-by-scale basis for the specific problem at hand.

Guo et al. [93] applied a speckle reduction method based on soft thresholding to logarithmically transformed SAR images and they found that a scale adaptive threshold between \( T = 1.5\sigma_n \) and \( T = 3\sigma_n \) yielded good results. Based on the Lorentz curve, Vidakovic [134] proposed a thresholding method in which the 100 \( \times \) \( p \)% of the wavelet coefficients with the smallest energy are set to zero, where \( p \) represents the proportion at which the gain by setting one more coefficient to zero will be smaller than the loss in the energy. Donoho and Johnstone [132] proposed the SureShrink method where the threshold is chosen on a scale-by-scale basis to minimise Stein’s unbiased risk estimate (SURE) [135]. Gagnon and Jouan [113] proposed a wavelet coefficient shrinkage filter based on symmetric Daubechies wavelets and found that it performs equally well as the classical adaptive filters for low-level noise and slightly better for high-level noise.

For all of these methods determination of the threshold requires a priori knowledge of or involves estimation of the noise variance. Some researchers approached the problem of threshold estimation by applying statistics techniques such as cross validation [136, 137] or generalised cross validation (GCV), independently proposed by Jansen and his colleagues [138, 139] and Weyrich and Warhola [140]. The advantage of these techniques is that they do not require prior knowledge of the noise variance. Weyrich and Warhola [140] also established the connection between the SureShrink method [132] and the GCV method. Application of Bayesian approach to analysis of the wavelet coefficients will be introduced later in the text (see page 102).

While hard thresholding is a “keep” or “kill” strategy, soft thresholding follows a “shrink” or “kill” rule. Despite this difference, hard and soft thresholding are similar in nature that they involve a “binary” decision process on the wavelet coefficients, which are deemed either due to noise and hence eliminated, or due to the sought-after real signal and hence kept or shrunk by the amplitude of the threshold. Zong [94, 141] proposed a nonlinear method for speckle reduction and contrast enhancement of echocardiographic images. Soft thresholding was only carried out on the wavelet coefficients of finer scales to suppress noise. Contrast enhancement was accomplished via a procedure called gen-
eralised adaptive gain followed by hard thresholding of the wavelet coefficients of middle scales, which is essentially a nonlinear operation on the wavelet coefficients, and specifically “tuned” for echocardiographic images [94]. Recently, Zhang [142] developed a new type of nonlinear thresholding functions for the proposed thresholding neural network (TNN) for adaptive noise reduction. Unlike the standard soft or hard thresholding function, the new thresholding functions are infinitely differentiable and used as the activation function of the TNN, which serves as a novel nonlinear adaptive filter.

Bayesian Wavelet Shrinkage

Recently there has been burgeoning interest in adopting a Bayesian approach to the analysis of the wavelet coefficients for noise reduction. These methods can be referred to as Bayesian wavelet shrinkage [98, 133, 143–149], which eventually lead to linear or nonlinear scaling of the wavelet coefficients (hence again the term shrinkage\(^1\)). The amount of shrinkage of the wavelet coefficients depend on their amplitudes. Generally, the larger the coefficients, the less they are shrunk. This greatly differs from the “binary” action of soft or hard thresholding (see Figure 5.1 on page 104).

Generally, these methods assume an \textit{a priori} probability model for the wavelet coefficients. For instance, by placing priors on the wavelet coefficients as a mixture of two normal distributions with different variances, Chipman, Kolaczyk, and McCulloch proposed the method of adaptive Bayesian wavelet shrinkage [133]. They approached the problem of denoising via a Bayesian framework and reduced it to estimation of \(w_s\) via the posterior mean \(E[w_s|w]\). It was shown this procedure is equivalent to the application of a nonlinear shrinkage function to the wavelet coefficients [133].

Chang, Yu, and Vetterli [146] proposed an adaptive Bayesian shrinkage approach based on modelling the prior for the wavelet coefficients as a generalised Gaussian (GG) distribution. Mih\textcuteck and his coworkers [150] developed a similar image denoising method, which is spatially adaptive as inspired by a wavelet image compression algorithm called the estimation-quantization (EQ) coder [151]. They modelled the wavelet coefficients as conditionally independent zero-mean Gaussian random variables with high local correlation, rather than GG random variables. Moreover, unlike [146], where the wavelet coefficients were modelled globally, they elected to exploit the \textit{local} properties (hence the spatial adaptability) of the wavelet coefficients in a small square-shaped window, and put a stochastic prior (a mixture of exponentials) on the local variances of them.

Simoncelli [145] adopted a slightly different prior, that is, a two-parameter generalised Laplacian density function (which encompasses the Gaussian density as a special case) of the following form

\[
p_{w_s}(w) = \frac{e^{-|\alpha w|^\beta}}{Z(\alpha, \beta)}
\]

\(^1\)Here “shrinkage” is a generalised term encompassing soft thresholding and other linear/nonlinear processing of the wavelet coefficients.
where \( Z(\alpha, \beta) = 2^\frac{n}{\beta} \Gamma(\frac{1}{\beta}) \) is a normalising constant. Assuming an additive white noise model, Simoncelli derived a standard maximum a posteriori (MAP) estimator, the special case of which has a linear form as follows when \( \beta = 2 \) (Gaussian density)

\[
\hat{w}_s = \frac{\sigma_s^2}{\sigma_s^2 + \sigma_n^2} w.
\]

Other values of \( \beta \) lead to a nonlinear form. Especially, \( \beta = 0.5 \) results in hard thresholding and \( \beta = 1 \) corresponds to soft thresholding \[145\]. Applying the minimum mean square error criterion, the author also demonstrated an optimal estimator of \( w_s \) similar to that derived by Chipman and his coworkers \[133\], that is, a Bayesian least-squares estimator provided by the mean of the posterior distribution. However, unlike \[133\], where a model of mixture of Gaussians is used, the Bayesian estimator in \[145\] generally does not have a closed-form expression except in the special Gaussian case when \( \beta = 2 \), which results in the same linear form as that of MAP estimator, given in Equation (5.26).

Recently, Achim, Bezerianos, and Tasakalides \[147–149\] proposed a Bayesian denoising algorithm based on an alpha-stable prior model for the wavelet coefficients. Like \[145\], the thresholding rules in \[147–149\] do not lead to a closed-form shrinking function, therefore numerical integration is required. The same situation happens in \[146\].

Vidakovic \[143\] also developed a threshold rule following a Bayesian approach, but from the different standpoint of hypothesis testing. The resultant Bayes factor thresholding procedure essentially resembles the hard thresholding \[152\]. Other Bayesian wavelet approaches can be found in the reviews given in \[98, 152, 153\] and the references therein. The latter two references \[152, 153\] are part of a collection of papers on Bayesian analysis of wavelet-based models, edited by Müller and Vidakovic \[154\], two active researchers in this area.

Figure 5.1 illustrates the concepts of wavelet coefficient hard thresholding (solid line) and soft thresholding (dashed line), and linear shrinkage (light colour solid line), which can result from a MAP estimator, and nonlinear shrinkage (light colour dashed line), which can result from Bayesian wavelet shrinkage.

Noise Discrimination by Singularity Detection

Motivated by the observation that the singularities of the signal have different properties from those of the white noise, Mallat and Hwang \[155, 156\] proposed a denoising method based on detection and analysis of singularities of a signal in the wavelet transform domain. Noise can be reduced by analysing the behaviour of the wavelet transform modulus maxima across different (dyadic) scales. The behaviour of the wavelet transform coefficients of a signal is determined by its local regularity, which can be measured with Lipschitz exponent \( \alpha \).

A function \( f(x) \) is referred to as Lipschitz \( \alpha \) if and only if the following hold for all \( x \)
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

Figure 5.1: Illustration of the concepts of hard (solid line) and soft (dashed line) thresholding, and linear (light colour solid line) and nonlinear (light colour dashed line) shrinkage of wavelet coefficients.

in a neighbourhood of $x_0$

$$|f(x) - f(x_0)| = K |x - x_0|^\alpha$$  \hspace{1cm} (5.27)

where $K$ is a constant. The Lipschitz regularity of a step function is 0. It was proved that the Lipschitz regularity of a signal $f(x)$ is related to its wavelet transform coefficients in that $f(x)$ is Lipschitz $\alpha$ if and only if the following holds

$$w_{2^j} = K(2^j)^\alpha$$  \hspace{1cm} (5.28)

where $w_{2^j}$ represents the wavelet transform coefficients of $f(x)$ at scale $2^j$. By finding the maximum exponent $\alpha$ that makes Equation (5.28) remain valid, their Lipschitz regularity can be computed and singularities be detected from the modulus maxima of the wavelet transform across scales.

The singularities of a signal are supposed to be no worse than discontinuity, that is, $\alpha > 0$. On the contrary, white Gaussian noise is almost singular everywhere with $\alpha < 0$. Therefore, as the scale increases, so does the amplitude of the wavelet transform coefficients of the signal. In contrast, the noise dominates the finest scale and gradually subsides as
the scale increases. Thus the noise can be separated from the signal by examining the behaviour of wavelet transform modulus maxima across different scales. Concretely, the local Lipschitz regularity can be estimated by searching for the largest exponent $\alpha$ that satisfies Equation (5.28) for a sequence of modulus maxima that persist across scales up to the finest scale. The modulus maxima due to the noise are not consistently of high amplitude, and decrease on average as scale increases, thus can be identified and removed.

Recent developments on noise reduction by singularity detection were reported in [157, 158], and [159].

Using Intra-scale or Inter-scale Dependency

Mallat and Hwang’s noise discrimination by singularity detection [155, 156] is essentially a method that takes advantage of inter-scale correlation of the wavelet coefficients. Similar ideas of employing inter-scale and intra-scale correlation of the wavelet coefficients for noise removal have motivated extensive research work in this area.

Xu and his colleagues [160] developed a noise filtration technique that relies on the multi-scale products of the wavelet coefficients at several adjacent scales, which serve as a “mask” to sift the image from noise. This method is inspired by the same observation that is behind Mallat and Hwang’s method in [155, 156], that is, the wavelet coefficients due to noise are only confined to a few fine scales whereas that of important features tend to propagate across different scales, especially those coarse ones.

Other methods based on inter-scale correlation include [161–163], and [144, 164, 165], which rely on the hidden Markov tree (HMT) model to capture inter-scale dependencies of the wavelet coefficients. Chipman [133] also suggested in his concluding remarks the possibility of employing inter-scale correlation as an extension to his Bayesian wavelet shrinkage method.

For modelling intra-scale dependencies of the wavelet coefficients [98, 166, 167] the Markov random field (MRF) model [168, 169] remains a very popular choice. This is also evidenced by the hybrid models such as [170], where both inter-scale and intra-scale correlation of the wavelet coefficients are explored. Other methods based on hybrid models include [145, 171].

5.2.3 Speckle Reduction by Geometry Driven Diffusion

Image processing and analysis methods based on partial differential equations (PDE’s) and geometry driven diffusion, especially as a promising tool for noise reduction, have received much attention in the last few decades. This is evidenced by the book edited by ter Haar Romeny [172] and more recently the “Special Issue on Partial Differential Equations and Geometry-Driven Diffusion in Image Processing and Analysis” of the IEEE Transactions on Image Processing in March 1998 [173]. Active contour models [174] and their variants or extensions such as Gradient Vector Flow (GVF) snakes [175, 176] (see Chapter 9 for
5.2. IMAGE LEVEL SPECKLE NOISE REDUCTION

details) all fall into this broad categories of methods in which PDE’s are involved in one way or another [173].

In these methods, an image is treated as a continuous function which deforms according to a partial differential evolution equation for which the original image is the initial condition. The solution of the PDE leads to a smoothed image. Since isotropic diffusion is equivalent to Gaussian smoothing, people are more interested in anisotropic diffusion, which was initiated by Perona and Malik [177]. To obtain a smoothed image, they proposed the following PDE

$$\frac{\partial I(x, y, t)}{\partial t} = \text{div}[c(||\nabla I||)\nabla I]$$

(5.29)

where \(\text{div}[\cdot]\) denotes the divergence operator and \(I(x, y, t)\) is an image in the continuous domain with \((x, y)\) denoting spatial coordinates and \(t\) an artificial time parameter. \(c(||\nabla I||)\) is referred to as “edge-stopping” function or diffusion coefficient function, where \(\nabla I\) is the image gradient and \(||\nabla I||\) is its magnitude.

When \(c(||\nabla I||)\) is a constant, Equation (5.29) reduces to the isotropic heat diffusion equation, which is equivalent to Gaussian filtering. Two nonlinear forms of “edge-stopping” function were employed in [177], namely,

$$c(||\nabla I||) = e^{-\frac{||\nabla I||^2}{K^2}},$$

(5.30)

which favours sharp edges over vague ones, and

$$c(||\nabla I||) = \frac{1}{1 + \frac{||\nabla I||^2}{K^2}},$$

(5.31)

which favours large smooth regions rather than smaller ones. \(K\) is a constant that characterises the magnitude of the gradients throughout the image [177]. Intuitively, \(c(||\nabla I||) \rightarrow 1\) when \(||\nabla I|| \ll K\), which means that in homogeneous areas isotropic diffusion (Gaussian smoothing) prevails. In contrast, if \(||\nabla I|| \gg K\), then \(c(||\nabla I||) \rightarrow 0\), that is, diffusion is “stopped” across edges where the gradients achieve the highest magnitude. This is the philosophy behind anisotropic diffusion for edge-sensitive smoothing of the images.

Recently, Black and his collaborators [178] explored the connections between anisotropic diffusion, robust statistics [179], and regularisation with line processes. They provided an interpretation of anisotropic diffusion from the perspective of robust statistics, and proposed a new “edge-stopping” function, which preserves sharp edges. This interpretation also leads naturally to edge detection since in the robust estimation framework edge points are treated as “outliers” between smoothed regions within the image. The authors also demonstrated the equivalence between anisotropic diffusion and regularisation with a line process for a particular class of robust error norms [178].

More recently, it was found that anisotropic diffusion methods are efficient for removing
additive noise but not the case for speckle noise, which is multiplicative and spatially correlated [72]. Yu and Acton [72] derived speckle reducing anisotropic diffusion (SRAD), specifically tailored for speckle removal in SAR and ultrasound images, by following the minimum mean square error (MMSE) criterion, which underpins the Lee [87, 88] and Frost [91] filters. The authors demonstrated the resemblance between the Lee and Frost filter and anisotropic diffusion and cast them into the framework of PDE to formulate a new adaptive edge-preserving SRAD method, which relies on the instantaneous coefficient of variation. The authors demonstrated that this method compares favourably with classical adaptive filtering methods and conventional anisotropic diffusion [72].

Unlike [72], where speckle is treated directly, Abd-Elmoniem and his colleagues considered logarithmically transformed images, where speckle noise approximately becomes additive white noise [101]. They proposed a nonlinear coherent diffusion (NCD) model for speckle reduction and coherence enhancement of ultrasound images. The dynamics of filtering, including the direction and magnitude of diffusion, is controlled by local coherence of the image which characterises regions with fully developed speckle, partially structured regions, and fully structured regions. The extent of speckle filtering varies between isotropic diffusion for fully formed speckle, anisotropic coherent diffusion for partially structured regions, and mean curvature motion [180] for fully structured regions.

5.2.4 Other Methods

Apart from the classical statistical filtering, wavelet methods, and geometry driven diffusion introduced above, other notable speckle reduction methods include:

- A maximum a posteriori (MAP) deconvolution algorithm for restoration of medical ultrasound images developed by Hokland and Kelly [181], which is based on Markov random field (MRF) models [168, 169]. First the reflectance of the tissue is modelled as the sum of a diffuse term (in which different tissues are modelled as independent Gaussian random variables) and an edge reflection term (which accounts for the large specular reflections that occur at some region boundaries). Then the observed image is modelled as the convolution of the tissue reflectance and the point spread function (PSF) of the imaging system, plus additive white Gaussian noise. A discrete MRF is employed to impose spatial smoothness constraints on different tissue regions and to segment these regions. Thus the problem of image restoration is reduced to deconvolution, which is accomplished by a prior estimate of the PSF followed by alternating steps of MAP estimation of tissue reflectance (restoration) and MAP estimation of the MRF labels of the tissue (segmentation). The authors demonstrated that the restoration results were better than that of Wiener filtering [181].

- Geometric filtering proposed by Crimmins [182–185], which is an iterative nonlinear algorithm based on the 8-hull algorithm and complementary hulling algorithm.

- The ultrasound image restoration method proposed by Schistad and Tant [186],
which includes the temporal filtering of consecutive frames for speckle reduction and spatial filtering of single frames for edge enhancement. This method is based on the quantum mechanical field model developed by Taut and Bølviken [187].

- A hybrid neural network proposed by Blacknell et al. [188] for speckle reduction of SAR images, which is based on a vector quantiser network and multi-layer perceptron (MLP) networks.

5.3 Discussion

In this chapter we have surveyed various methods for speckle reduction, which we classified into system level methods (such as compound scanning and Hokland’s harmonic imaging) and post-image-formation filtering techniques. Although we deemed that pre-image-formation speckle reduction techniques involving imaging system internals are promising and warrant further research, we have since been concentrating on off-line image processing techniques, which we loosely categorize into

- classical adaptive filters, which are based on local statistics of the images;
- wavelet methods, most of which are inspired by Donoho’s wavelet shrinkage;
- anisotropic diffusion, the underpinning framework of which is partial differential equations; and
- other methods such as geometric filtering and neural networks.

Most of the noise reduction methods we have seen so far are designed with an additive white noise model in mind, although sometimes implicitly. It is commonly accepted that speckle is multiplicative, however, this is not a problem since it can be circumvented by logarithmically transforming the images before application of any noise reduction methods. Some researchers directly dealt with logarithmically compressed images [27, 40].

It is a common practice that noise reduction is formulated as an estimation problem where the aim is to recover the original image from its noisy observation by following some criteria such as minimum mean square error (MMSE) or maximum a posteriori (MAP). Examples of imposing the former criterion include the well known classical adaptive filters such as the Lee [87, 88], Frost [91] and Kuan [90] filter. The latter criterion leads to MAP filters such as [104], extended by Lopez and his colleagues [126, 127]. It is not surprising to see that the same idea is prevalent among the most up-to-date filtering methods in the wavelet domain, such as the work of Argenti [97] (MMSE) and Foucher [96] (MAP).

The wavelet transform offers a scale-space localisation framework for analysis of an image, which results in partial decorrelation of signal and noise in that the noise energy concentrates in the few fine scales, and quickly subsides as the scale increases, whereas the signal energy dominates the coarse scales and the scaling coefficients. Thus the problem
of denoising reduces to design of a function that “shrinks” the wavelet coefficients deemed to be due to noise. By far the most influential work on denoising by wavelets is Donoho’s wavelet shrinkage [99, 100, 132], which has inspired extensive research activity in this area, especially Bayesian wavelet shrinkage [154], into which the classical Bayesian framework is naturally incorporated for analysing the wavelet coefficients.

While in most of the wavelet methods the wavelet coefficients are assumed to be independently distributed, the exploration of inter-scale and intra-scale dependency of the wavelet coefficients is burgeoning [154]. Many researchers resort to Markov random field (MRF) models [168, 169] or to hidden Markov trees for capturing this correlation [98, 144, 164–167]. Our speckle reduction scheme to be described in Chapter 6 is inspired by wavelet shrinkage in general and the work of Scharcanski [163] in particular, but specifically adapted for ultrasound images based on statistical modelling of the wavelet coefficients.

Gagnon and Jouan [113] compared their complex Wavelet Coefficient Shrinkage (WCS) filter, which is an elliptical soft thresholding procedure, with some classical adaptive filters, and found that performance of their wavelet method was no worse than that of the classical speckle filters in comparison. Many researchers who advocate wavelet methods, from simple soft thresholding to non-trivial Bayesian wavelet shrinkage, draw similar conclusions. Generally speaking, we deem that compared to classical adaptive filters, wavelet methods have advantage in that the wavelet transform opens a new horizon via multi-scale representation of an image, thus providing more possibility and flexibility in analysis and treatment of speckle noise, such as exploration of inter-scale and intra-scale dependency, which is not possible or readily available with the classical adaptive methods. However, classical speckle filters have their own advantage in that most of them are well established, especially in the SAR imaging community, due to their simplicity and fast implementation.

Incorporation of a segmentation procedure into the speckle reduction process, either explicitly or implicitly, is a promising paradigm that is expected to find wide application. For example, Hokland’s unsupervised deconvolution algorithm [181] consists of alternating steps of restoration and segmentation, which is based on MRF modelling. In [101] segmentation is implicit and determined by speckle extent and image anisotropy. Karaman’s adaptive speckle suppression technique [108] features a smoothing kernel of varying shape and size, which depends on local statistics based region growing, essentially a segmentation process. In [98] a MRF model is used to differentiate the wavelet coefficients between significant (edge) and non-significant (non-edge) ones. Another example of noise reduction on the basis of segmentation is [105]. All these practice indicate that segmentation and denoising are not necessarily two seemingly irrelevant areas, but rather can be integrated into the speckle reduction process.

Another example that two seemingly different areas at first sight can be combined is simultaneous noise reduction and compression (quantisation) [146, 189–194]. The idea of filtering noise via data compression can be dated back at least to [189, 192]. These methods exploit the fact that random noise does not have redundancies thus is difficult
to compress, whereas a signal contains structural features for which a concise representation with reduced entropy can be obtained via the quantisation process of a compression algorithm. Concretely, it has been shown that quantisation of wavelet coefficients with a “dead-zone” around zero is approximately equivalent to soft thresholding [146, 194].
Chapter 6

Speckle Reduction Using the Dyadic Wavelet Transform

In this chapter we propose a noise reduction scheme based on Mallat and Zhong’s dyadic wavelet transform \[195\]. The results of applying this scheme to both simulated B-scan ultrasound images and real ultrasound images of the prostate are promising. The important features of the original image are preserved while most of the speckle noise is removed.

6.1 Introduction

In Chapter 4 we have seen that the Rayleigh distribution is adequate for modelling speckle in B-scan ultrasound images of a homogeneous phantom. However, as the phantom structure becomes complex with more objects being introduced, it is virtually impossible to find a single “all-around” distribution model to describe all possible scenarios of speckled ultrasound images. Nevertheless, computer simulation of B-scan ultrasound images provides a viable test-bed for validation of speckle reduction algorithms. While our computer simulation only models a simplified version of a real ultrasound imaging system, we do not think it is appropriate to degrade a real B-mode ultrasound image by introducing “extra” multiplicative random noise as in \[147\]. A real ultrasound image is already a speckled image simply because speckle is an inherent part of the imaging process. There is no point in introducing “aliens” while there are “home-growns”.

As we have seen in Chapter 5 many speckle reduction techniques have been developed in recent years based on various speckle distribution models. Among them the classical adaptive filters are the most established methods, which make use of local statistics and/or have control parameters, the optimum values of which are difficult to find in practice. While some of them achieve superior performance in smoothing speckle in homogeneous areas, others are better in preserving sharp local structures such as edges. In this chapter we present a speckle reduction scheme based on statistical modelling of the wavelet coefficients. It is intended to address these two areas of concern at the same time.

A brief introduction to wavelets and multiresolution analysis in general, and the
6.2 Wavelets and the Dyadic Wavelet Transform

Wavelets and wavelet analysis are one of the most exciting developments in engineering and mathematics during the last two decades. The advent of wavelets makes an excellent example of how important and beneficial the interaction between mathematicians, physicists, and engineers can be. In particular, the last ten years have seen an explosion of applications of wavelets in such diverse fields as signal processing, subband coding and data compression, computer graphics, biomedical engineering, and many other areas. This is evidenced by the special issues on wavelets in the Proceedings of the IEEE [196], which contains a collection of overview tutorials on wavelets and applications, and the IEEE Transactions such as [197–200].

There are many classical books on wavelets as well. Meyer [201] and Daubechies [202] both gave a thorough treatment of the mathematical aspects of wavelets. Vetterli and Kovačević [203] provide expositions of wavelets in the context of subband coding, whereas Mallat [204, 205] discusses wavelets from the viewpoint of signal processing. An accessible introduction was provided by Strang and Nguyen [206], and Burrus et al. [207]. Rioul and Vetterli [208] also gave an excellent tutorial on wavelet and signal processing, and provided pointers on history of wavelets in the extended references. See [209] for an informal review of the history and philosophy behind wavelets and wavelet analysis.

6.2.1 Wavelets and the Continuous Wavelet Transform

So what are wavelets anyway? Simply speaking, wavelets are families of scaled and shifted versions of a kernel function \( \psi(t) \) (referred to as the “mother” wavelet)

\[
\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t - b}{a} \right), \quad a > 0, \ b \in \mathbb{R}
\]

where \( a \) is a scaling factor, which “compresses” or “dilates” the wavelets, and \( b \) represents the translation factor. A mother wavelet must be admissible, satisfying

\[
C_\psi = \int_{-\infty}^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega < +\infty
\]

where \( C_\psi \) is referred to as the admissibility constant, which is determined by the mother wavelet \( \psi(t) \). \( \Psi(\omega) = \int_{-\infty}^{+\infty} \psi(t)e^{-i\omega t} dt \) is the Fourier transform of \( \psi(t) \). Equation (6.2)
holds only if \( \Psi(0) = 0 \), which implies 
\[
\int_{-\infty}^{+\infty} \psi(t)dt = \Psi(0) = 0 \quad [210].
\]
Therefore, a mother wavelet oscillates with an average value of zero and usually features an irregular and asymmetric waveform of limited duration.

The role of wavelets \( \psi_{a,b}(t) \) is analogous to that of the functions \( e^{i\omega t} \) with regard to the Fourier transform, that is, they are the building blocks of the wavelet transform. Mathematically, given a mother wavelet \( \psi(t) \), the continuous wavelet transform (CWT) of a function \( f \in L^2(\mathbb{R}) \) is defined by

\[
Wf(a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{ \sqrt{a}} \psi^* \left( \frac{t-b}{a} \right) dt \quad (6.5)
\]

The admissibility of the mother wavelet \( \psi(t) \) ensures the existence of an inverse transform and \( f(t) \) can be recovered from \( Wf(a,b) \) by

\[
f(t) = \frac{1}{ C_\psi} \int_0^{+\infty} \frac{da}{a^2} \int_{-\infty}^{+\infty} Wf(a,b) \psi_{a,b}(t) db. \quad (6.6)
\]

In other words, the truncated integral

\[
\int_{-\frac{A}{a}}^{\frac{B}{a}} Wf(a,b) \psi_{a,b}(t) db
\]

converges to \( C_\psi f(t) \) as \( A, B \to +\infty \) [210].

Wavelets are a subject that is neither very new nor very old [211]. It is not very new because the underlying idea of breaking a function into building blocks or representing it with linear expansions of basis functions dates back to early 19th century when Joseph Fourier introduced the transform that bears his name. Moreover, Alfred Haar constructed the first wavelet in 1910, although it was not referred to as wavelet at that time. On the other hand, wavelet became a term with special meaning from 1980s, especially after wavelets and wavelet analysis were formalised and advocated [201, 202].

Just as the Fourier transform breaks up a signal into sinusoids of various frequencies, amplitudes and phases, the wavelet transform represents a signal by the family of shifted and scaled (compressed or expanded) versions of the mother wavelet. However, there are several nontrivial differences between the basis functions of a Fourier transform and a

\[
L^2(\mathbb{R}) \text{ represents a Hilbert space, that is, the space of square-integrable (Lebesgue-integrable) complex-valued functions, namely, functions } f \text{ satisfying } ||f||^2 = \int_{-\infty}^{+\infty} |f(t)|^2 dt < +\infty \quad [203]. \text{ The scalar product of function } f(t) \text{ and } g(t) \text{ on } L^2(\mathbb{R}) \text{ is given by}
\]

\[
\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)g^*(t)dt \quad (6.3)
\]

where \( g^*(t) \) represents the complex conjugate of \( g(t) \). The norm of \( f(t) \) is given by

\[
||f|| = \sqrt{\langle f, f \rangle} = \sqrt{\int_{-\infty}^{+\infty} |f(t)|^2 dt}. \quad (6.4)
\]

Note in the following text we assume that \( \psi(t) \) is a real function, thus the \( ^* \) in Equation (6.3) and (6.5) can be removed. Some researchers have found complex wavelets to be useful as well [113].
Firstly, unlike a sinusoid which oscillates with equal amplitude from \(-\infty\) to \(+\infty\), a wavelet decays very quickly to zero with limited duration;

- While a sinusoid is smooth and predictable, a wavelet usually features an irregular and asymmetric shape;

- A wavelet has an average value of zero and finite energy because of its localised support, whereas a sinusoid has infinite energy.

Because of these discrepancies, a wavelet transform differs from a Fourier transform in many ways. Fourier analysis provides us with a perspective in the frequency domain for a signal. One drawback of the Fourier transform is that the time information of a signal is totally lost after it is transformed to the frequency domain. This problem is partially addressed by conventional time-frequency signal representations such as the short-time Fourier transform (STFT), which, however, suffers from a trade-off problem of fixed time resolution versus frequency resolution. Moreover, Fourier techniques are not efficient in dealing with a discontinuity or an abrupt change in a signal. Recall the classical example that it takes many sinusoids of different frequencies to cancel one another to represent the sharp corner of a rectangular wave producing the Gibbs ringing phenomenon.

In contrast, wavelet analysis produces a time-scale view or multi-resolution representation of a signal. Intuitively, low scales correspond to compressed wavelets, which reveal rapidly changing details, namely, the high frequency components of a signal. High scales correspond to stretched wavelets, which represent coarse level signal transitions. Roughly speaking, the scaling parameter \(a\) in Equation (6.5) localises analysis of a signal \(f\) in frequency and the shift parameter \(b\) localises analysis in time. In other words, the wavelet transform leads to simultaneous time-frequency localisation in depicting the information contained in a signal. Moreover, some wavelets have sharp transitions themselves. Therefore, one major advantage offered by wavelets is the ability to perform local analysis, that is, the wavelet’s localised support means that some local features, such as abrupt transitions of a signal, can be better depicted or revealed by wavelets.

There is another notable difference between the Fourier transform and wavelet transform. In Fourier-based techniques, the underlying (trigonometric) basis functions are unique. In contrast, it does not make sense to advocate a wavelet transform without referring to the associated mother wavelet, since there are actually many “flavours” of mother wavelets. The simplest among them is the Haar wavelet, which is now understood as a special case of the Daubechies wavelets [202]. In this thesis, we focus on the quadratic spline wavelet advocated by Mallat and Zhong [195].

Despite the advantages of the wavelet transform over the traditional Fourier transform, wavelets and wavelet analysis should not be hailed as a “miracle” tool, as warned by Kovačević and Daubechies [211]. Wavelet analysis should only be regarded as an alternative or supplement to Fourier analysis, not a replacement. The concepts of wavelets
and wavelet analysis cannot be elaborated without reference to Fourier analysis, and some wavelets were actually designed from the ground up in the frequency domain.

### 6.2.2 Multiresolution Analysis

Before the advent of the wavelet theory, many researchers had recognised the necessity of adopting a coarse-to-fine multiresolution approach for image analysis. These practices can be traced back to Marr [212], Rosenfeld and Vanderberg’s multi-scale template matching [213], Witkin’s scale-space filtering [214, 215], and Burt and Adelson’s Laplacian pyramid [216]. Mallat [217] and Meyer [201], formalised and refined these ideas in the framework of wavelets and multiresolution analysis.

Simply speaking, multiresolution analysis is about approximation, namely, approximation in the framework of wavelets. In a multiresolution analysis a signal is recursively decomposed into approximations plus details at different resolutions. The approximations are the projections of the signal onto the subspaces (of $L^2(\mathbb{R})$) spanned by some functions $\{\phi_{j,k} \mid j, k \in \mathbb{Z}\}$ at different scales, which are defined as the dilated and shifted versions of what is referred to as the scaling function $\phi(t)$,

$$
\phi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \phi\left(\frac{t}{2^j} - k\right)
$$

where $\phi(t)$ satisfies $\int_{-\infty}^{\infty} \phi(t)dt = 1$.

The details, namely the information that is lost between two successive approximations from one scale to another coarser scale, belong to the subspaces (of $L^2(\mathbb{R})$) spanned by the wavelets $\{\psi_{j,k} \mid j, k \in \mathbb{Z}\}$, where

$$
\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi\left(\frac{t}{2^j} - k\right)
$$

are dilated and shifted versions of the mother wavelet $\psi(t)$, which satisfies $\int_{-\infty}^{\infty} \psi(t)dt = 0$.

### Concepts of Multiresolution Analysis and the Scaling Function

Formally, a multiresolution analysis of a signal $f(t) \in L^2(\mathbb{R})$ is the projection of $f(t)$ onto a sequence of embedded closed subspaces $\{V_j \mid j \in \mathbb{Z}\}$, which satisfy the following requirements [201–203, 206, 207, 217]:

1. Nesting of spaces\(^2\)

$$
\cdots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots,
$$

which for simplicity can be written as

$$
V_j \subset V_{j-1}, \quad \text{for all } j \in \mathbb{Z}.
$$

\(^2\)Note here we adopt the same nesting order of the subspaces as that in [202, 203], that is, the smaller the index, the larger the space. Some authors adopt a reverse ordering [206, 207].
2. Completeness: including upward completeness

\[ \bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R}) \]  

(6.10)

where \( \bigcup \) represents union and the over-line represents closure\(^3\) of the subspaces, and downward completeness

\[ \bigcap_{j \in \mathbb{Z}} V_j = \{0\} \]  

(6.11)

where \( \bigcap \) denotes intersection and \( \{0\} \) represents the null space. Equations (6.9)-(6.11) imply

\[ \lim_{j \to -\infty} V_j = L^2(\mathbb{R}) \]  

(6.12)

\[ \lim_{j \to +\infty} V_j = \{0\}. \]  

(6.13)

3. Scale invariance

\[ f(t) \in V_0 \iff f(2^{-j}t) \in V_j, \]  

(6.14)

in other words, all the spaces are scaled version of the “central space” \( V_0 \).

4. Shift invariance under integer translations

\[ f(t) \in V_0 \iff f(t - k) \in V_0, \text{ for all } k \in \mathbb{Z}. \]  

(6.15)

5. Existence of a basis: there exists a scaling function \( \phi \in V_0 \), such that \( \{\phi(t-k) \mid k \in \mathbb{Z}\} \) form an orthonormal basis\(^4\) for \( V_0 \).

There are many spaces that satisfy Equations (6.9)-(6.11) but they may not necessarily realise a multiresolution analysis [202]. The “heart and soul” of a multiresolution analysis lies in Equations (6.14)-(6.15) plus, of course, the last requirement.

From the last requirement and Equation (6.14), it immediately follows that the functions \( \{\phi_{j,k} \mid j, k \in \mathbb{Z}\} \) (see Equation (6.7)) form an orthonormal basis for \( V_j \) for all \( j \in \mathbb{Z} \). Thus any signal \( f(t) \in L^2(\mathbb{R}) \) can be approximated by its orthogonal projection onto \( V_j \) as

\[ \tilde{f}_j(t) = \sum_{k \in \mathbb{Z}} a_{j,k} \phi_{j,k}(t) \]  

(6.17)

\(^3\)The closure of the subspaces contains not only all signals that can be represented by a linear expansion of the basis functions of the subspaces, but also the ones that are the limits of these expansions.

\(^4\)For the sake of simplicity we assume here the orthonormality of \( \{\phi(t-k) \mid k \in \mathbb{Z}\} \), namely

\[ \langle \phi(t-k), \phi(t-l) \rangle = \int_{-\infty}^{+\infty} \phi(t-k) \phi(t-l)dt = \delta(k-l), \quad \text{for all } k, l \in \mathbb{Z}. \]  

(6.16)

However, this restriction is not necessary and can be relaxed (such that \( \{\phi(t-k) \mid k \in \mathbb{Z}\} \) constitute a Riesz basis for \( V_0 \)) since they can always be orthogonalised [202,203]. For these cases, dual scaling functions and wavelets are needed for synthesis.
where

$$a_{j,k} = (f(t), \phi_{j,k}(t)) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^j}} \phi \left( \frac{t}{2^j} - k \right) dt.$$  \hspace{1cm} (6.18)

Equation (6.10), which means that the union of all the subspaces $V_j$ is dense in $L^2(\mathbb{R})$, ensures that

$$\lim_{j \to -\infty} \widetilde{f}_j = f. \hspace{1cm} (6.19)$$

In other words, such a projection can give an arbitrary close approximation of the signal $f(t)$ as the scale decreases.

### Wavelets

In practice, however, it is intractable to achieve arbitrary close approximations by decreasing the scale (increasing the size of the subspaces $V_j$), or we are more interested in the details that are needed to reconstruct the approximation at scale $2^{j-1}$ from that at scale $2^j$. These details might reveal important information about the signal that is not readily available in its original form (or disguise). That is where the wavelets $\{\psi_{j,k} \mid j, k \in \mathbb{Z}\}$ (see Equation (6.8)) come into play.

Before we define the wavelets, let us first introduce a new sequence of subspaces $\{W_j \mid j \in \mathbb{Z}\}$ of $L^2(\mathbb{R})$, which are defined as the orthogonal complement\(^5\) of $V_j$ in $V_{j-1}$, namely $V_j \perp W_j$ and

$$V_{j-1} = V_j \oplus W_j \hspace{1cm} (6.22)$$

where $\oplus$ stands for direct sum\(^6\) of two subspaces. Since $W_j \perp V_j$ and $W_{j+1} \subset V_j$, it immediately follows that

$$W_j \perp W_{j+1} \text{ for all } j \in \mathbb{Z}, \hspace{1cm} (6.23)$$

namely, the subspaces $\{W_j \mid j \in \mathbb{Z}\}$ are mutually orthogonal\(^7\) (note unlike $V_j$, they are not nested, only mutually orthogonal).

Then starting from subspace $V_0$ and recursively applying Equation (6.22) we can

---

\(^5\)This means that if $\{\psi_{j,k} \mid k \in \mathbb{Z}\}$ constitute an orthonormal basis for $W_j$, namely,

$$\langle \psi_{j,k}(t), \psi_{j,l}(t) \rangle = \int_{-\infty}^{+\infty} \psi_{j,k}(t) \psi_{j,l}(t) dt = \delta(k-l), \text{ for all } j, k, l \in \mathbb{Z}, \hspace{1cm} (6.20)$$

then the following holds

$$\langle \phi_{j,k}(t), \psi_{j,l}(t) \rangle = \int_{-\infty}^{+\infty} \phi_{j,k}(t) \psi_{j,l}(t) dt = 0, \text{ for all } j, k, l \in \mathbb{Z}. \hspace{1cm} (6.21)$$

\(^6\)In general direct sums are not necessarily orthogonal sums. For this case they are. For biorthogonal wavelets they are not.

\(^7\)Namely the following holds

$$\langle \psi_{j,k}(t), \psi_{j',k'}(t) \rangle = \int_{-\infty}^{+\infty} \psi_{j,k}(t) \psi_{j',k'}(t) dt = \delta(j-j') \delta(k-k'), \text{ for all } j, k, j', k' \in \mathbb{Z}. \hspace{1cm} (6.24)$$
6.2. WAVELETS AND THE DYADIC WAVELET TRANSFORM

Figure 6.1: Illustration of the concepts of multiresolution analysis and the relationship between the subspaces of $L^2(\mathbb{R})$ (adapted from [207]). $W_j$ (subspaces spanned by wavelets) are orthogonal complements of $V_j$ (spanned by scaling functions): $V_j \perp W_j$. The $V_j$ are nesting $V_j \subset V_{j-1}$, and the $W_j$ are mutually orthogonal $W_j \perp W_{j-1}$ for all $j \in \mathbb{Z}$.

Rewrite Equation (6.10) as

$$L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_{-1} \oplus W_{-2} \oplus \cdots$$

(6.25)

which can be generalised as

$$L^2(\mathbb{R}) = V_J \oplus W_J \oplus W_{J-1} \oplus W_{J-2} \oplus \cdots$$

$$= V_J \oplus \left( \bigoplus_{j \leq J} W_j \right)$$

(6.26)

when we arbitrarily choose an initial subspace $V_J$. When $J \to +\infty$, it follows from Equation (6.10) and (6.11) that Equation (6.26) can be rewritten as

$$L^2(\mathbb{R}) = \cdots \oplus W_2 \oplus W_1 \oplus W_0 \oplus W_{-1} \oplus W_{-2} \oplus \cdots$$

$$= \bigoplus_{j \in \mathbb{Z}} W_j$$

(6.27)

which means that $L^2(\mathbb{R})$ can be decomposed into the mutually orthogonal subspaces $\{W_j | j \in \mathbb{Z}\}$. Figure 6.1 illustrates the concepts of multiresolution analysis and the relationship between $V_j$ and $W_j$, the subspaces of $L^2(\mathbb{R})$.

Now we define a function $\psi(t) \in W_0$, namely the mother wavelet, such that the $\{\psi(t - \cdots$
$k \mid k \in \mathbb{Z}$ constitute an orthonormal basis for $W_0$. Since the $W_j \subset V_{j-1}$ and the $V_{j-1}$ satisfy the scale invariance property in Equation (6.14), it automatically follows that the $W_j$ inherit the same property, namely

$$f(t) \in W_0 \iff f(2^{-j}t) \in W_j. \quad (6.28)$$

Therefore the wavelets $\{\psi_{j,k} \mid k \in \mathbb{Z}\}$ (Equation (6.8)) form an orthonormal basis for subspace $W_j$, for any $j \in \mathbb{Z}$. Thus, due to Equations (6.10) and (6.11) the whole collection $\{\psi_{j,k} \mid j, k \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R})$.

Thus from Equation (6.27) we can see that any function $f(t) \in L^2(\mathbb{R})$ can be represented as an expansion of these basis functions, namely

$$f(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f(t), \psi_{j,k}(t) \rangle \psi_{j,k}(t) \quad (6.29)$$

where

$$d_{j,k} = \langle f(t), \psi_{j,k}(t) \rangle = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^j}} \psi \left( \frac{t}{2^j} - k \right) dt \quad (6.30)$$

are the discrete wavelet transform coefficients of $f(t)$ at scale $2^j$. Equation (6.30) and (6.29) respectively establish the discrete wavelet transform and its inverse. As we can see, by virtue of the (forward) transform (Equation (6.30)) a function $f(t)$ is decomposed into the wavelet coefficients $d_{j,k}$, from which the function is reconstructed through the inverse transform (Equation (6.29)).

Equation (6.29) has the same problem as Equation (6.17) in that it is intractable in practice. Therefore a viable approach is to represent a function $f(t)$ as an expansion by virtue of both the scaling functions $\phi_{j,k}(t)$ and the wavelets $\psi_{j,k}(t)$. Intuitively, the wavelets $\{\psi_{j,k} \mid k \in \mathbb{Z}\}$ at scale $2^j$ characterise the disjoint differences between two immediately nesting subspaces $V_j$ and $V_{j-1}$, namely, the “details” that is needed to reconstruct the approximation at scale $2^{j-1}$ from that at scale $2^j$. Thus from Equation (6.22) and (6.17) it follows

$$\tilde{f}_{j-1}(t) = \tilde{f}_j(t) + \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t) = \sum_{k \in \mathbb{Z}} a_{j,k} \phi_{j,k}(t) + \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t). \quad (6.31)$$

The implication of the above equation is that starting from the “approximation” of the signal $\tilde{f}_j(t)$ at an arbitrary coarse scale $2^J$ and repeatedly adding “details” at subsequent scales we can obtain an approximation of the signal at any resolution. Thus from
Equation (6.26) it follows
\[
f(t) = \tilde{f}(t) + \sum_{j=-\infty}^{J} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t)
\]
\[
= \sum_{k \in \mathbb{Z}} a_{J,k} \phi_{J,k}(t) + \sum_{j=-\infty}^{J} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t). \quad (6.32)
\]
In practice, however, one is usually given a sampled version of the real signal which is measured at a finite resolution (or sampling rate). Therefore one cannot realise a multiresolution analysis of the signal at an arbitrary fine resolution. The finest scale one can achieve is at the sample level. Without loss of generality we can normalise the finest scale to 1 to model this resolution limitation. Thus Equation (6.32) becomes
\[
f(t) = \sum_{k \in \mathbb{Z}} a_{J,k} \phi_{J,k}(t) + \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} d_{j,k} \psi_{j,k}(t)
\]
\[
= \sum_{k \in \mathbb{Z}} \langle f(t), \phi_{J,k}(t) \rangle \phi_{J,k}(t) + \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \langle f(t), \psi_{j,k}(t) \rangle \psi_{j,k}(t) \quad (6.33)
\]
(from Equation (6.18) and (6.30)). Therefore, starting from any coarse scale \(2^J\), the sequence of discrete coefficients\(^8\)
\[
\{a_{J,k}, d_{j,k} | 1 \leq j \leq J\}
\]
together constitute a discrete wavelet transform (DWT) of the signal \(f(t)\), where the coarse level scaling coefficients \(a_{J,k} = \langle f(t), \phi_{J,k}(t) \rangle\) characterise the signal components above scale \(2^J\).

Filter Banks

Now the task of realising a multiresolution analysis reduces to contriving a scaling function \(\phi \in V_0\) such that \(\{\phi(t-k) | k \in \mathbb{Z}\}\) forms an orthonormal basis for \(V_0\), and a mother wavelet \(\psi \in W_0\) such that \(\{\psi(t-k) | k \in \mathbb{Z}\}\) constitutes an orthonormal basis for \(W_0\). Due to the nesting of subspaces (Equation (6.9)) and the scaling property (Equation (6.14)), we have \(\phi(t) \in V_0 \subset V_{-1}\). Since \(\{\phi_{-1,k}(t) = \sqrt{2} \phi(2t - k) | k \in \mathbb{Z}\}\) is an orthonormal basis for \(V_{-1}\), we can express \(\phi(t)\) as a linear expansion of these basis function as
\[
\phi(t) = \sum_{k=-\infty}^{+\infty} h[k] \sqrt{2} \phi(2t - k) \quad (6.34)
\]
\(^8\)Here we rewrite the coefficients \(a_{J,k}\) and \(d_{j,k}\) to discriminate between the scale factor \(j\) and the discrete time (or translation) index \(k\).
where \( h[k] = \langle \phi(t), \phi_{-1,k}(t) \rangle = \int_{-\infty}^{+\infty} \phi(t) \sqrt{2} \phi(2t - k) dt \). Equation (6.34) is referred to as the refinement equation, dilation equation, or multiresolution analysis equation. Taking the Fourier transform of both sides of this equation, we have

\[
\Phi(\omega) = \frac{1}{\sqrt{2}} H \left( \frac{\omega}{2} \right) \Phi \left( \frac{\omega}{2} \right)
\]

(6.35)

where the Fourier transform of \( \phi(t) \) is \( \Phi(\omega) = \int_{-\infty}^{+\infty} \phi(t) e^{-i\omega t} dt \) and \( i = \sqrt{-1} \). The discrete-time Fourier transform (DTFT) of \( h[k] \) is

\[
H(\omega) = \sum_{k \in \mathbb{Z}} h[k] e^{-i\omega k}
\]

(6.36)

which is a \( 2\pi \)-periodic function in \( L^2([0,2\pi]) \). It can be shown that it characterises a multiresolution analysis. Especially, \( H(\omega) \) can be viewed as the DTFT of a discrete-time finite impulse response (FIR) filter \( h[k] \), which establishes the connection between discrete and continuous time, namely, continuous-time wavelet bases can be constructed from discrete iterated filters. Recursively applying Equation (6.34) we obtain

\[
\Phi(\omega) = \Phi(0) \prod_{j=1}^{+\infty} \frac{H(2^{-j}\omega)}{\sqrt{2}}
\]

(6.37)

since \( \lim_{j \to +\infty} \Phi(2^{-j}\omega) = \Phi(0) = \int_{-\infty}^{+\infty} \phi(t) dt = 1 \).

Equation (6.37) is the basis for theoretically proving the existence and uniqueness (corresponding to a set of \( h[k] \)) of the scaling function \( \phi(t) \), and for obtaining the waveform of \( \phi(t) \) in a form of successive approximations (which is sometimes known as the cascade algorithm), although the waveform of \( \phi(t) \) (and that of mother wavelet \( \psi(t) \)) is seldom explicitly needed in practice.

Similarly, since \( \psi(t) \in W_0 \subset V_{-1} \), we can represent \( \psi(t) \) as a linear expansion of \( \{ \phi_{-1,k}(t) = \sqrt{2} \phi(2t - k) | k \in \mathbb{Z} \} \), namely

\[
\psi(t) = \sum_{k=-\infty}^{+\infty} g[k] \phi(2t - k)
\]

(6.38)

where \( g[k] = \langle \psi(t), \phi_{-1,k}(t) \rangle = \int_{-\infty}^{+\infty} \psi(t) \sqrt{2} \phi(2t - k) dt \). After taking the Fourier transform of both sides of Equation (6.38) we obtain

\[
\Psi(\omega) = \frac{1}{\sqrt{2}} G \left( \frac{\omega}{2} \right) \Phi \left( \frac{\omega}{2} \right)
\]

(6.39)
6.2. WAVELETS AND THE DYADIC WAVELET TRANSFORM

where the discrete-time Fourier transform of $g[k]$ is

$$G(\omega) = \sum_{k \in \mathbb{Z}} g[k] e^{-i\omega k} \tag{6.40}$$

and the Fourier transform of $\psi(t)$ is $\Psi(\omega) = \int_{-\infty}^{+\infty} \psi(t) e^{-i\omega t} dt$. Analogous to $h[k]$ and $H(\omega)$, $g[k]$ and $G(\omega)$ can be respectively regarded as a FIR filter and its frequency response.

It can be shown that the DWT of a signal $f$, namely the scaling coefficients $a_j$ and wavelet coefficients $d_j$, can be obtained recursively through convolving the signal with low-pass analysis (decomposition) filter $h[-k]$ and band-pass analysis filter $g[-k]$ followed by downsampling (keeping every even term) [201–203, 206, 207, 217]

$$a_j(k) = \sum_n h[n - 2k] a_{j-1}(n) \tag{6.41}$$

$$d_j(k) = \sum_n g[n - 2k] a_{j-1}(n) \tag{6.42}$$

with $a_0 = f$ being the original discretised signal. For reconstruction of the original signal $f$ from its DWT we have

$$a_{j-1}(k) = \sum_n h[k - 2n] a_j(n) + \sum_n g[k - 2n] d_j(n) \tag{6.43}$$

where $a_j$ and $d_j$ are upsamled (inserting a zero between each pair of terms) before being
filtered respectively by the low-pass synthesis (reconstruction) filter $h[k]$ and the band-pass synthesis filter $g[k]$. It can be shown that for a multiresolution analysis that involves an orthonormal wavelet system, the analysis filters are the time-reversed version of the synthesis filters, as we have seen above. Moreover, for an orthonormal wavelet system there is a special relationship between the two filters, namely, $g[k] = (-1)^k h[1 - k]$. Figure 6.2 illustrates the filtering process for a two-stage multiresolution analysis, where $\downarrow 2$ and $\uparrow 2$ respectively denote downsampling and upsampling.

Thus a multiresolution analysis is characterised by an orthonormal wavelet basis, which gives rise to a pair of discrete-time filters $h[k]$ (low-pass) and $g[k]$ (band-pass), or their counterparts in the frequency domain $H(\omega)$ and $G(\omega)$, which have appropriate properties in the time and frequency domain. These filters not only characterise the waveforms of the scaling function and wavelets, but also allow the implementation of fast algorithms for computation of the discrete wavelet transform as a filtering process. In most practical applications, the scaling function and wavelets are seldom used explicitly for signal analysis and synthesis, but rather the associated filter banks. Thus the link between a multiresolution analysis and filter banks is established.

### 6.2.3 Biorthogonal Wavelets

So far we have been concentrating on orthogonal wavelet systems. For a multiresolution analysis involving orthogonal wavelets and scaling functions, the same set of scaling functions $\{\phi_{j,k} \mid j, k \in \mathbb{Z}\}$ and wavelets $\{\psi_{j,k} \mid j, k \in \mathbb{Z}\}$ are used for both analysis and synthesis, as we have seen in Equations (6.17)-(6.18), (6.29)-(6.30), and (6.33). Although orthogonal wavelet systems are mathematically elegant, the requirement of the orthogonality places a strong constraint on construction of possible wavelet systems. For example, the associated filters must be of the same even length, and the scaling function and wavelets are asymmetric.

If the orthogonality requirement is relaxed, greater flexibilities can be achieved in designing and implementing a multiresolution analysis. For instance, at the cost of dropping orthogonality, biorthogonal wavelet systems [218, 219] allow symmetry and antisymmetry of the scaling functions and wavelets, which are suitable for image processing, but not available in the framework of orthogonal wavelets.

In a biorthogonal wavelet system there are two sets of multiresolution subspaces involved, one for analysis and a dual one for synthesis. The multiresolution subspaces for analysis are similar to those corresponding to an orthogonal wavelet system except that the orthogonality requirement is relaxed, namely, the subspaces $V_j$ are not perpendicular to $W_j$ anymore. In other words, they are complements of one and another in $V_{j-1}$, but unlike that in Equation (6.22) the direct sum is not an orthogonal one. Apart from this set of subspaces, there exists another set of subspaces $\tilde{V}_j$ and $\tilde{W}_j$, which satisfy

$$\tilde{V}_j \subset \tilde{V}_{j-1}, \quad \text{for all } j \in \mathbb{Z},$$

(6.44)
Figure 6.3: Illustration of the relationship between the subspaces of a biorthogonal wavelets system (adapted from [206]). $W_j$ is not orthogonal to $V_j$ but to its dual $\tilde{V}_j$, and $\tilde{W}_j$ is not orthogonal to $\tilde{V}_j$ but to its dual $V_j$.

\[ \tilde{V}_{j-1} = \tilde{V}_j \oplus \tilde{W}_j, \quad \text{for all } j \in \mathbb{Z} \]

(6.45)

(note that as for $V_j$ and $W_j$ the direct sum here is not orthogonal), such that the following hold

$W_j \perp \tilde{V}_j$ and $\tilde{W}_j \perp V_j$, \quad \text{for all } j \in \mathbb{Z},

(6.46)

which means that $W_j$ is not orthogonal to $V_j$ but to its dual $\tilde{V}_j$, and $\tilde{W}_j$ is not orthogonal to $\tilde{V}_j$ but to its dual $V_j$. Figure 6.3 illustrates the relationship between the subspaces of a biorthogonal wavelet system.

The dual scaling functions $\{\tilde{\phi}_{j,k} \mid j, k \in \mathbb{Z}\}$ form bases for the subspaces $\tilde{V}_j$ and the dual wavelets $\{\tilde{\psi}_{j,k} \mid j, k \in \mathbb{Z}\}$ constitute bases for $\tilde{W}_j$. Thus Equation (6.46) implies

\[ \langle \phi(t - k), \tilde{\phi}(t - l) \rangle = \int_{-\infty}^{+\infty} \phi(t - k) \tilde{\phi}(t - l) dt = \delta(k - l), \quad \text{for all } k, l \in \mathbb{Z}, \]

(6.47)

\[ \langle \psi_{j,k}(t), \tilde{\psi}_{j,l}(t) \rangle = \int_{-\infty}^{+\infty} \psi_{j,k}(t) \tilde{\psi}_{j,l}(t) dt = \delta(k - l), \quad \text{for all } j, k, l \in \mathbb{Z}, \]

(6.48)

\[ \langle \phi_{j,k}(t), \tilde{\psi}_{j,l}(t) \rangle = \int_{-\infty}^{+\infty} \phi_{j,k}(t) \tilde{\psi}_{j,l}(t) dt = 0, \quad \text{for all } j, k, l \in \mathbb{Z}, \]

(6.49)

\[ \langle \psi_{j,k}(t), \tilde{\psi}_{j',k'}(t) \rangle = \int_{-\infty}^{+\infty} \psi_{j,k}(t) \tilde{\psi}_{j',k'}(t) dt = \delta(j - j') \delta(k - k') \quad \text{for all } j, k, j', k' \in \mathbb{Z}, \]

(6.50)

which mean that the scaling functions $\phi(t)$ are orthogonal to their dual $\tilde{\phi}(t)$ and similarly
the wavelets $\psi(t)$ are orthogonal to the dual wavelet $\tilde{\psi}(t)$ (hence the term “biorthogonal”). Equations (6.47)-(6.50) respectively are comparable to the orthogonality requirements of an orthogonal wavelet system in Equations (6.16), (6.20), (6.21), and (6.24). Accordingly the counterparts of Equations (6.29) and (6.33) are respectively

$$ f(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f(t), \psi_{j,k}(t) \rangle \tilde{\psi}_{j,k}(t) $$ (6.51)

and

$$ f(t) = \sum_{k \in \mathbb{Z}} \langle f(t), \phi_{J,k}(t) \rangle \tilde{\phi}_{J,k}(t) + \sum_{j=1}^{J} \sum_{k \in \mathbb{Z}} \langle f(t), \psi_{j,k}(t) \rangle \tilde{\psi}_{j,k}(t), $$ (6.52)

in which dual wavelets and scaling functions are used for synthesis.

We have seen in the previous subsection that the scaling functions and wavelets of an orthogonal wavelet system satisfy the dilation equations (6.34) and (6.38), which lead to the design of FIR filters for computation of a fast wavelet transform (see Equations (6.34)-(6.43)). In a biorthogonal wavelet system, the same equation applies for the scaling functions and wavelets for analysis and those for synthesis, which lead to FIR filters for decomposition $h[k]$ and $g[k]$ and dual filters $\tilde{h}[k]$ and $\tilde{g}[k]$ for reconstruction.

In order to achieve perfect reconstruction of the original signal, it can be shown that the analysis and synthesis wavelets must satisfy the following requirement in the frequency domain [204, 205]

$$ \sum_{j=-\infty}^{\infty} \Psi^*(2^j \omega) \tilde{\Psi}(2^j \omega) = 1, $$ (6.53)

where $\Psi(\omega)$ and $\tilde{\Psi}(\omega)$ are respectively the Fourier transform of the analysis wavelet $\psi(t)$ and synthesis wavelet $\tilde{\psi}(t)$. It can be shown [204, 205] that this is equivalent to

$$ H^*(\omega)\tilde{H}(\omega) + G^*(\omega)\tilde{G}(\omega) = 2, \quad \text{for all } \omega \in [-\pi, \pi], $$ (6.54)

where $H(\omega)$, $G(\omega)$, $\tilde{H}(\omega)$, and $\tilde{G}(\omega)$ respectively denote the DTFT of the corresponding FIR filters for decomposition $h[k]$ and $g[k]$ and dual filters $\tilde{h}[k]$ and $\tilde{g}[k]$ for reconstruction.

To obtain the DWT of a signal, for decomposition Equations (6.41)-(6.42) still apply for a biorthogonal wavelet system. However, for reconstruction Equation (6.43) now becomes

$$ a_{j-1}(k) = \sum_n \tilde{h}[k - 2n] a_j(n) + \sum_n \tilde{g}[k - 2n] d_j(n). $$ (6.55)

### 6.2.4 Mallat and Zhong’s Dyadic Wavelet Transform

Although wavelets have found wide applications in many areas such as signal processing, subband coding and data compression, they have been less useful in the field of visual pattern recognition because the traditional wavelet transform based on an orthogonal wavelet system is not translation invariant [220]. Mallat and Zhong [195] advocated a
6.2. WAVELETS AND THE DYADIC WAVELET TRANSFORM

The dyadic wavelet transform (DWT) which is translation invariant and oriented for multiscale edge detection. This transform features quadratic spline wavelets, and the corresponding scaling functions closely approximate a Gaussian function. Shown in Figure 6.4 are the waveforms of the quadratic spline wavelet and the associated scaling function.

The fast algorithm (which is called algorithme à trous in French [204, 205, 221]) that implements the dyadic wavelet transform is similar to that of a biorthogonal wavelet transform, but without downsampling, which makes translation invariance possible. As for a biorthogonal wavelet system, the FIR filters corresponding to the scaling functions $\phi$ and wavelets $\psi$ and their duals $\tilde{\phi}, \tilde{\psi}$ are respectively denoted as $h, g, \tilde{h}, \tilde{g}$. The Fourier transform of the scaling function $\phi(t)$ (a box spline of degree 2) is given by

$$\Phi(\omega) = \left(\sin\frac{\omega}{2}\right)^3 e^{-i\frac{\omega}{2}}. \tag{6.56}$$

From Equation (6.35) we have

$$H(\omega) = \sqrt{2} \frac{\Phi(2\omega)}{\Phi(\omega)} = \sqrt{2} \left(\cos\frac{\omega}{2}\right)^3 e^{-i\frac{\omega}{2}}. \tag{6.57}$$

The wavelet $\psi(t)$ is constructed such that it has one vanishing moment\(^9\) and the corresponding $G(\omega)$ satisfying this constraint is given by

$$G(\omega) = -i\sqrt{2} \left(\sin\frac{\omega}{2}\right) e^{-i\frac{\omega}{2}}. \tag{6.58}$$

\(^9\)The moments of a wavelet $\psi(t)$ are defined as $m_\nu = \int_{-\infty}^{\infty} x^\nu \psi(t) dt$. A function is referred to as having $N$ vanishing moments if $m_\nu=0$, for $0 \leq \nu < N$, and $m_N \neq 0$ [202].
SPECKLE REDUCTION USING THE DYADIC WAVELET TRANSFORM

Table 6.1: Coefficients of the FIR filters associated with the quadratic spline wavelets and the scaling functions. \( h[k] \) and \( g[k] \) respectively represents the low-pass and band-pass filters for decomposition. \( \tilde{h}[k] \) and \( \tilde{g}[k] \) are the dual filters for reconstruction. \( \tilde{l}[k] \) is the filter required for reconstruction of images.

<table>
<thead>
<tr>
<th>( k )</th>
<th>( h[k] )</th>
<th>( g[k] )</th>
<th>( \tilde{h}[k] )</th>
<th>( \tilde{g}[k] )</th>
<th>( \tilde{l}[k] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>-3</td>
<td></td>
<td></td>
<td></td>
<td>0.0078125</td>
<td></td>
</tr>
<tr>
<td>-2</td>
<td></td>
<td></td>
<td>-0.03125</td>
<td>0.046875</td>
<td></td>
</tr>
<tr>
<td>-1</td>
<td>0.125</td>
<td>0.125</td>
<td>-0.21875</td>
<td>0.1171875</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.375</td>
<td>-0.5</td>
<td>-0.6875</td>
<td>0.65625</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.375</td>
<td>0.5</td>
<td>0.6875</td>
<td>0.1171875</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.125</td>
<td>0.125</td>
<td>0.21875</td>
<td>0.046875</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>0.03125</td>
<td>0.0078125</td>
<td></td>
</tr>
</tbody>
</table>

Table 6.1 lists the coefficients of the FIR filters corresponding to the quadratic spline scaling functions and wavelets defined in Equations (6.56)-(6.60). As we can see, \( h[k] \) is a low-pass filter similar to a moving average filter, smoothing or blurring the signal. \( g[k] \) is a band-pass filter, comparable to an edge detector, locating abrupt changes of the signal as local extrema. \( \tilde{l}[k] \) is the filter required for dealing with reconstruction of two-dimensional signals (images), and its DTFT is determined by

\[
\tilde{L}(\omega) = \frac{1 + |H(\omega)|^2}{2}.
\]

Refer to [195, 204, 205] for more details on the design of the quadratic spline scaling functions, wavelets and the associated FIR filters. An alternative set of FIR filters can be found in [94] and [222].

Then the DWT of a signal can be obtained by following a fast cascading filtering algorithm similar to that illustrated in Figure 6.2 and Equations (6.41)-(6.42), but without downsampling of the transform coefficients. Concretely, by recursively convolving a signal \( f \) with the decomposition filters \( h[k] \) (low-pass) and \( g[k] \) (band-pass) associated with the quadratic spline wavelet and its scaling function, we decompose the signal into “approximations” \( a_j \) (scaling coefficients) and “details” \( d_j \) (wavelet coefficients) at different
6.2. WAVELETS AND THE DYADIC WAVELET TRANSFORM

\[
a_j[k] = a_{j-1}[k] \ast h_{j-1}[-k],
\]

\[
d_j[k] = a_{j-1}[k] \ast g_{j-1}[-k]
\]

(6.62)

(6.63)

where \( \ast \) denotes circular convolution and \( a_0 = f \) represents the original signal (of finite length). \( h_j[k] \) and \( g_j[k] \) stands for the discrete filters at each scale \( 2^j \), which are obtained by inserting \( (2^j - 1) \) zeros, thus creating "holes" (trous in French) between each pair of coefficients of \( h[k] \) and \( g[k] \) respectively. It can be shown that their DTFTs are respectively \( H(2^j \omega) \) and \( G(2^j \omega) \). \( h_j[-k] \) and \( g_j[-k] \) are the time-reversed versions of \( h_j[k] \) and \( g_j[k] \).

As we have seen in section 6.2.2 (see page 120), in practice the highest resolution (the lowest scale) we can achieve of a signal is that of the sampled signal itself. It can be shown that for Mallat and Zhong’s DWT, as scales increases, \( a_j \) ultimately reduces to the average value of the original signal \( a_0 = f \). Thus the scales available for us to explore are limited to \( 1 \leq j \leq J \), where \( 2^J \) is the maximum decomposition scale dependent on \( N \), the length of the original signal, \( J = \log_2 N \). In practical applications, however, it is seldom necessary to decompose a signal up to the maximum scale \( 2^J \). Usually a scale \( 2^{J'} \) slightly lower than \( 2^J \) is enough. Thus the coefficients

\[
\{a_{j'}[k], d_{j}[k] \mid 1 \leq j \leq J'\}
\]

form a dyadic wavelet transform (DWT) of the signal, where the coarse level scaling coefficients \( a_{J'}[k] \) characterise the signal components above scale \( 2^{J'} \).

For reconstruction, the dual low-pass filter \( \tilde{h}[k] \) and band-pass filter \( \tilde{g}[k] \) are used, and the original signal \( a_0 \) can be recovered from its DWT through the following cascading algorithm which implements the inverse DWT

\[
a_j[k] = \frac{1}{2} \left( a_{j+1}[k] \ast \tilde{h}_j[k] + d_{j+1}[k] \ast \tilde{g}_j[k] \right)
\]

(6.64)

where \( \tilde{h}_j[k] \) and \( \tilde{g}_j[k] \) are similarly obtained from \( \tilde{h}[k] \) and \( \tilde{g}[k] \) by inserting zeros.

Since this wavelet system is nonorthogonal, the details \( \{d_j\}_{1 \leq j \leq J} \) plus the coarse information \( a_J \) give an over-complete representation of the signal in the wavelet transform domain. Moreover, what distinguishes this transform from a conventional discrete wavelet transform based on an orthogonal wavelet system is that there is no decimation (down-sampling) for decomposition at each scale. It is this information redundancy that makes it translation invariant, which is a desired property for computer vision, pattern recognition and related areas. This redundancy implies that prominent features of the signal, such as sharp edges, which are presented as local extrema in the “details” \( \{d_j\} \) of the wavelet transform, tend to propagate correlatively across many scales, and most importantly, align at approximately the same location. In contrast, the noise manifests itself in a random pattern and fades out quickly as the scale increases.

This feature can be demonstrated in Figure 6.5, in which a synthetic signal and its
Figure 6.5: A synthetic signal corrupted with noise and its DWT. (a) The synthetic signal $a_0$, with three noise corrupted box functions respectively located between 50–80, 110–160, and 180–206. (b)-(f) Details of the signal at five scales, $d_1$-$d_5$. (g) Approximation of the signal $a_5$ at scale $2^5$.

discrete wavelet transform are shown. The synthetic signal is composed of three box functions (respectively located between 50–80, 110–160, and 180–206) corrupted by noise, as shown in Figure 6.5(a), where the edges of the box functions are indicated by vertical dotted lines. The signal is decomposed into five levels up to scale $2^5$, and the details are shown in Figure 6.5(b)-(f). Figure 6.5(g) shows the approximation, namely a smoothed version of the signal at scale $2^5$.

As we can see, the edges are presented as local extrema in the DWT details $d_1$-$d_5$ and across the different scales aligned at the same locations. Noise is only dominant at the lower scales. In contrast, as the scale increases, so does the strength of the local extrema corresponding to the edges. Note how the first descending edge located at 80 gains “strength” at higher scales despite the fact that it has no sheer advantage over the noise at lower scales. This is one of the most attractive characteristics of this transform and is the underlying basis of the edge based noise reduction scheme that takes advantage
of multiscale correlation, which will be introduced in the next section.

Extending this fast algorithm to the two-dimensional case, the dyadic wavelet transform of an image $a_0$ is obtained by filtering row-wise and column-wise separately by the (one-dimensional) band-pass filter $g[k]$ to give the horizontal “details”

$$d^H_j[m,n] = a_{j-1}[m,n] \star (g_{j-1}[-m], \delta[n])$$ (6.65)

and vertical “details”

$$d^V_j[m,n] = a_{j-1}[m,n] \star (\delta[m], g_{j-1}[-n])$$ (6.66)

where $m, n$ respectively represent the row and column indices of the image pixels. $I \star (X, Y)$ denotes separable convolution of the image $I$ with two one-dimensional filters, row-wise with filter $X$, and column-wise with filter $Y$. $\delta[m]$ denotes the discrete Dirac function\(^\text{10}\).

Because the band-pass filter $g[k]$ is essentially an “edge detector”, at each scale horizontal “details” $d^H_j$ characterise vertical edges and vertical “details” $d^V_j$ characterise horizontal edges. Thus for each pixel $d^H_j[m,n]$ and $d^V_j[m,n]$ together constitute a gradient vector

\(^{10}\text{The discrete Dirac function is defined as } \delta[m - m'] = \begin{cases} 1 & \text{if } m - m' = 0 \\ 0 & \text{otherwise} \end{cases} \text{ for all } m, m' \in \mathbb{Z}.\)
Figure 6.7: The DWT details of the simulated image shown in Figure 6.6(a): horizontal details $d_j^H$ (left column) and vertical details $d_j^V$ (right column), from top to bottom $j = 1, 2, 3, 4$. 
and the modulus of the gradient vector is given by

\[ M_j[m, n] = \sqrt{\left| d^H_j[m, n] \right|^2 + \left| d^V_j[m, n] \right|^2}, \] (6.67)

which are multiscale edge maps of the original image.

At each scale, \( a_{j-1} \) (with \( a_0 \) being the original image) is also filtered both row-wise and column-wise consecutively by the low-pass filter \( h[k] \) to give the “approximations” at the next scale

\[ a_j[m, n] = a_{j-1}[m, n] \star (h_{j-1}[-m], h_{j-1}[-n]). \] (6.68)

In summary, at each level \( a_{j-1} \) is decomposed into three components (of the same size as that of the original image), namely, horizontal details \( d^H_j \), vertical details \( d^V_j \), and approximations \( a_j \), which will be successively filtered to give another set of three component images at the next level. Moreover, \( d^H_j \) and \( d^V_j \) are combined to give modulus \( M_j \).

For reconstruction the dual filters \( \tilde{h}[k], \tilde{g}[k], \) and \( \tilde{l}[k] \) are used respectively to filter \( a_{j+1}, d^H_{j+1}, \) and \( d^V_{j+1} \), and the filtered outputs are combined to give \( a_j \)

\[
\begin{align*}
    a_j[m, n] &= a_{j+1}[m, n] \star (\tilde{h}_j[m], \tilde{h}_j[n]) \\
    &+ d^H_{j+1}[m, n] \star (\tilde{g}_j[m], \tilde{l}_j[n]) \\
    &+ d^V_{j+1}[m, n] \star (\tilde{l}_j[m], \tilde{g}_j[n]).
\end{align*}
\] (6.69)

Refer to [195, 204, 205] for more details on implementation of the fast dyadic wavelet transform.

As an example, Figure 6.6(a) shows a simulated B-scan ultrasound image, which contains seven discs of different sizes. The object/background contrast is set to 6.9dB. The image is decomposed up to scale \( 2^4 \) by the fast DWT algorithm described above. The DWT approximation at scale \( 2^4 \), \( a_4 \), is shown in Figure 6.6(b). As we can see, although \( a_4 \) is noise-free and presents rough structures of the objects, it is also over-smoothed and has lost the details that are needed to form sharp edges.

The details lost at each scale are shown in Figure 6.7, where the left column illustrates the horizontal details \( d^H_j \) and the right column shows the vertical details \( d^V_j \) (from top to bottom \( j = 1, 2, 3, 4 \)). Note the pixel values of each component image are normalised, with black, gray, and white pixels respectively representing negative, near zero, and positive wavelet coefficients. (The component images are also scaled down in size to fit on one page.) Figure 6.8 shows the modulus, or multiscale edge maps at each scale obtained from \( d^H_j \) and \( d^V_j \) according to Equation (6.67).

As we can see in Figures 6.7 and 6.8, the edges of the objects are emerging from “behind” the noise as scale increases. The lower scales are dominated by noise, as expected.

---

\(^{11}\)Equation (6.67) is the same as Equation (10.7) (page 212) in Chapter 10 except minor notation changes. In Chapter 10 subscript \( 2^j \) is used to emphasise the dyadic scales, whereas \( j \) is used for this purpose in this chapter. This will not cause any confusion given the context.
The multiscale edge maps in Figure 6.8 are comparable to what we can obtained from applying a Canny edge detector [223, 224] or a LoG operator (see Chapter 7 and 8), where the standard deviation of the Gaussian kernel plays the role of varying degree of smoothing, comparable to scale variation in a DWT. However, either a Canny edge detector or a LoG operator only provides a means for signal analysis, none for signal synthesis. Namely, no definitive means is available to combine the edge maps at different scales to construct a noise-free signal. Moreover, compromise has to be made in choosing the standard deviation of the Gaussian kernel between feature preservation and noise reduction.

In contrast, Mallat and Zhong’s DWT provides a mathematically-sound framework
not only for signal analysis (which we have employed to design the contour extraction scheme integrating snakes in Chapter 10), but also for signal synthesis, which makes it possible to somehow manipulate the wavelet coefficients at different scales to reconstruct a noise-free signal from the modified wavelet coefficients. This will be the topic of next section, in which we introduce a noise reduction scheme in the wavelet transform domain.

6.3 Noise Reduction Scheme in the Wavelet Transform Domain

As we have seen in the previous section, the discrete wavelet transform, especially Mallat and Zhong’s dyadic wavelet transform [195], provides a way of systematically decomposing a signal into approximations and details and most importantly, reconstructing the original signal from them. While the approximations closely follow the gross “structure” of the signal, the details contain mostly fine components or high frequency features, including abrupt signal changes and of course, noise. Many noise reduction methods have been proposed based on the wavelet transform. Most of them involve the manipulation of the details (or wavelet coefficients) leaving the approximations (or scaling coefficients) intact as noise is mostly contained in the details, especially those at lower scales. For a review of these methods, especially wavelet shrinkage and Bayesian wavelet shrinkage, refer to subsection 5.2.2 and the references therein.

Although noise is mostly confined to the lower scales, as we have seen in Figures 6.5, 6.7, and 6.8, we cannot completely remove the details of lower scales as they also contain important high frequency information, such as that of the edges. Somehow we need a way to “sift” out the wavelet coefficients due to true edges while leaving behind the responses of the noise and eliminate them to recover a noise-free signal with legitimate fine features preserved. As we have seen in Figures 6.5, and 6.8, the edges are presented as local extrema of the wavelet coefficients and most importantly are correlated in a way that they are aligned at the same locations across the different scales. This feature can be exploited to reinforce the local extrema due to real abrupt transitions of the signal and suppress those due to noise as they tend to subside as the scale increases.

For instance, Xu [160] proposed a method for noise removal in magnetic resonance images by exploiting multi-scale products of the wavelet coefficients of the different scales. Similar ideas can be traced back to Rosenfeld and co-workers [225, 226] before the advent of the wavelet framework. In [160] only additive noise is considered. Moreover, effective removal of the noise depends on manual selection of a threshold, which requires preselection of a relatively “signal-free” area. (In other words, a priori knowledge about the image structure and noise is required to some extent.) Refer to subsection 5.2.2 and the references therein for other methods that exploit intra-scale or inter-scale correlation of the wavelet coefficients.

In this section to alleviate the effects of speckle in ultrasound images we propose a
noise reduction scheme in the wavelet transform domain as inspired by [163], which is on the basis of modelling statistical properties of the wavelet coefficients, and assumes no a priori knowledge about the image structure.

### 6.3.1 Modelling Statistical Properties of the Wavelet Coefficients

Many probability models have been proposed to characterise the statistical properties of the wavelet coefficients, including generalised Gaussian (GG) [146,165,167,217], conditionally independent Gaussian [150], generalised Laplacian [145,227], alpha-stable [147–149], and two-component Gaussian mixture [98,133,144,165]. Refer to [228] for a review of some of the distribution models for the wavelet coefficients.

Every model mentioned above has its own justifications, but on the other hand is not an exact model suitable to any image as well. Although the Gaussian mixture model may not be as accurate as GG model for some images, we found it simple, efficient, intuitive, and most importantly, realistic enough for our applications, which are attractive features also noted by other researchers [144,165].

The two-component Gaussian mixture model, for its name’s sake, characterises the probability density function of the wavelet coefficients (of either component image $d_j^H$ or $d_j^V$) as a mixture of two normal distributions with zero mean and different variances, namely,

$$ p_{d_j}(x) = \sum_i w_i G_{\sigma_i}(x), \quad i \in \{0, 1\} \quad (6.70) $$

where

$$ G_{\sigma_i}(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-x^2/2\sigma_i^2}, \quad i \in \{0, 1\} \quad (6.71) $$

are two Gaussian functions with mean zero and variance $\sigma_i^2$, where $\sigma_0^2 < \sigma_1^2$. $w_i$ are the corresponding “weights” of the two normal distributions, satisfying $w_0 + w_1 = 1$.

Formally, an $N$-state Gaussian mixture model for a random variable $X$ consists of [144]

- a discrete random state variable $S$, which takes the values $s = i$, $i \in \{0, 1, \cdots, N-1\}$ according to a priori probability $P(s = i)$, with $\sum_i P(s = i) = 1$; and

- the corresponding conditional probability density functions (pdf) $p_{X|S}(x|s = i)$, which are defined as zero-mean Gaussian functions with variance $\sigma_i^2$

$$ p_{X|S}(x|s = i) = G_{\sigma_i}(x) = \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-x^2/2\sigma_i^2}, \quad i \in \{0, 1, \cdots, N-1\}. \quad (6.72) $$

Thus the overall pdf of the random variable $X$ is given by

$$ p_X(x) = \sum_{i=0}^{N-1} P(s = i) p_{X|S}(x|s = i), \quad (6.73) $$

which is generally non-Gaussian due to the randomness of state variable $S$. Equation (6.73)
6.3. NOISE REDUCTION SCHEME IN THE WAVELET TRANSFORM DOMAIN

Figure 6.9: Modelling the distribution of the wavelet coefficients as a two-state zero-mean Gaussian mixture. (a) Histogram (light colour solid line) of the wavelet coefficients of $d^{H}_3$ and the fitted Gaussian mixture (dark colour solid line). The two Gaussian components are respectively denoted by the dotted (low variance) and dashed line (high variance). (b) Those of $d^{L}_3$. 

136
SPECKLE REDUCTION USING THE DYADIC WAVELET TRANSFORM

can be interpreted as follows. An observation of the random variable \( X \) is obtained by first drawing a state value \( s = i \) according to \( P(s = i) \), followed by drawing a sample \( x \) according to the corresponding conditional pdf \( p_{X|S}(x|s = i) = G_{\sigma_i}(x) \), which is a Gaussian.

Although as noted by some researchers [144], for any density with a finite number of discontinuities better fit can be achieved by increasing the number of states and/or allowing non-zero means, the two-component zero mean Gaussian mixture model remains a sensible choice because it is robust, manageable, and accurate enough for our purpose. Thus Equation (6.73) reduces to

\[
p_d(x) = P(s = 0) p_{d|S}(x|s = 0) + P(s = 1) p_{d|S}(x|s = 1) = w_0 \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-x^2/2\sigma_0^2} + w_1 \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-x^2/2\sigma_1^2}
\]  

(6.74)

for a two-component zero mean Gaussian mixture model. Each wavelet coefficient falls in either state 0 with probability \( P(s = 0) = w_0 \), or state 1 with probability \( P(s = 1) = 1 - P(s = 0) = 1 - w_0 = w_1 \). The wavelet coefficients of the former “category” can be interpreted as “insignificant” ones of low variance \( \sigma_0^2 \), which are mainly due to noise. Those of the latter “category” correspond to “significant” ones of high variance \( \sigma_1^2 \), which are mostly due to meaningful features such as prominent object boundaries.

Figures 6.9 and 6.10 respectively demonstrate a two-state zero-mean Gaussian mixture model fitted to the histograms of the wavelet coefficients \( d_{3}\ ), \( d_{3}\ ) at scale \( 2^3 \) and \( d_{4}\ ), \( d_{4}\ ) at scale \( 2^4 \) (shown in Figure 6.7). Histograms of the wavelet coefficients are denoted by light colour solid lines and the fitted Gaussian mixtures are presented as dark colour solid lines. The two component Gaussians are respectively denoted by dotted and dashed lines. The hyperparameters \( \sigma_0, \sigma_1, \) and \( w_0 \) can be estimated by the methods introduced in section 3.2 such as the method of CDF (fitting the cumulative distribution function of a model to the empirical cumulative distribution function). Refer to that section for more details.

As we can see, the two-state Gaussian mixture model fits very well to the histograms of the wavelet coefficients. Moreover, as the scale increases from \( 2^3 \) to \( 2^4 \), so does the “weight” \( w_1 \), namely, the probability that a wavelet coefficient is “significant”. This agrees with our previous observation that meaningful object features are more pronounced at higher scales than at lower scales, which are dominated by noise. This in turn implies that the Gaussian mixture model we have adopted is sensible.

Thus following the form of Equation (6.74) and considering Equation (6.67), it is acceptable to approximately model the modulus \( M_j \) as a mixture of two Rayleigh distributions, namely,

\[
p_{M_j}(x) = P(s = 0) p_{M_j|S}(x|s = 0) + P(s = 1) p_{M_j|S}(x|s = 1) = w_0 \frac{x}{\sigma_0^2} e^{-x^2/2\sigma_0^2} + w_1 \frac{x}{\sigma_1^2} e^{-x^2/2\sigma_1^2}
\]  

(6.75)
6.3. **NOISE REDUCTION SCHEME IN THE WAVELET TRANSFORM DOMAIN**

![Graph](image_url)

**Figure 6.10:** Modelling the distribution of the wavelet coefficients as a two-state zero-mean Gaussian mixture. (a) Histogram (light colour solid line) of the wavelet coefficients of $d_H^4$ and the fitted Gaussian mixture (dark colour solid line). The two Gaussian components are respectively denoted by the dotted and dashed lines. (b) Those of $d_V^4$. 

138
where the conditional probability density functions $p_{M_j|S}(x|s = i), i \in \{0, 1\}$ are defined as a Rayleigh function (see Equation (3.2)).

Figure 6.11 illustrates a two-state Rayleigh mixture model fitted to the histograms of the modulus of component wavelet coefficients $M_3$ at scale $2^3$ and $M_4$ at scale $2^4$ (shown in Figure 6.8). Shown in Figure 6.11(a) are the histogram of modulus $M_3$, obtained from $d_{H}^M$ (the histogram of which is shown in Figure 6.9(a)) and $d_{V}^M$ (histogram shown in Figure 6.9(b)) according to Equation (6.67), the two component Rayleighs and their mixture. Figure 6.11(b) shows those of modulus $M_4$.

Note in Figure 6.11(b) how the mode of the high-variance Rayleigh component shifts further to the right of that of the low-variance one, while in Figure 6.11(a) it is still overshadowed by its counterpart of low-variance. Moreover, the probability that a wavelet coefficient is “significant”, which is indicated by the “weight” $w_1$, increases from 0.3217 at scale $2^3$ to 0.3317 at scale $2^4$, as expected. All of these observations indicate that the noise (insignificant wavelet coefficients or non-edges) and real signal (significant wavelet coefficients or edges) are being decorrelated, and most importantly, the latter are more pronounced as scale increases.

Apart from modelling the modulus $M_j$ as a mixture of two Rayleigh distributions, which is the practice in [163], we can also elect to model the squared modulus, namely

\[ M_j^2[m, n] = |d_{H}^M[m, n]|^2 + |d_{V}^M[m, n]|^2 \]  

(6.76)

as a two-component mixture of Gamma distributions\(^\text{12}\)

\[ p_{M_j^2}(x) = P(s = 0) p_{M_j^2|S}(x|s = 0) + P(s = 1) p_{M_j^2|S}(x|s = 1) \]

\[ = w_0 \frac{1}{\lambda_0} e^{-x/\lambda_0} + w_1 \frac{1}{\lambda_1} e^{-x/\lambda_1} \]  

(6.78)

where the conditional probability density functions $p_{M_j^2|S}(x|s = i)$ are governed by a special case Gamma function with parameter $\alpha = 1$ (see Equation (6.77) or [49])

\[ p_{M_j^2|S}(x|s = i) = \frac{1}{\lambda_i} e^{-x/\lambda_i}, \quad \lambda_i > 0, \quad i \in \{0, 1\} \]  

(6.79)

Figure 6.12 shows a two-state Gamma mixture model fitted to the histograms of the squared modulus of wavelet coefficients $M_3^2$ (Figure 6.12(a)) at scale $2^3$ and $M_4^2$ (Figure 6.12(b)) at scale $2^4$. As we can see, the two-state Gamma mixture model is a reasonable model for characterising statistical properties of the squared modulus of wavelet

\(^\text{12}\)The density function of a random variable $X$ following a Gamma distribution is given by

\[ p_X(x) = \frac{1}{\lambda^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\lambda} \]  

(6.77)

where $x \geq 0$ and $\alpha, \lambda > 0$ [49]. The sum of $k$ mutually independent, squared Gaussian random variables (with zero-mean and unit-variance) follows a chi-squared density function with $k$ degrees of freedom. A chi-squared random variable is a special case of Gamma random variable [49, 229].
6.3. NOISE REDUCTION SCHEME IN THE WAVELET TRANSFORM DOMAIN

Figure 6.11: Modelling the distribution of the modulus of component wavelet coefficients as a two-state Rayleigh mixture. (a) Histogram (light colour solid line) of the modulus $M_3$ and the fitted Rayleigh mixture (dark colour solid line). The two Rayleigh components are respectively denoted by the dotted and dashed lines. (b) Those of $M_4$. 

140
Figure 6.12: Modelling the distribution of the squared modulus of wavelet coefficients as a two-state Gamma mixture. (a) Histogram (light colour solid line) of the squared modulus \( M_2^2 \) and the fitted Gamma mixture (dark colour solid line). The two Gamma components are respectively denoted by the dotted and dashed lines. (b) Those of \( M_4^2 \) (horizontal axis \( \times 10^{-3} \)).
coefficients $M_j^2$. As the scale increases, a wavelet coefficient due to real features is more likely to be “significant”, as indicated by the increase of “weight” $w_1$ from 0.3378 to 0.3443. This again agrees with our previous observations.

In summary, so far we have been concentrating on modelling the statistical properties of the wavelet coefficients. We model the component wavelet coefficients $d_j^H$ and $d_j^V$ as a two-state zero-mean Gaussian mixture, the modulus $M_j$ as a two-state Rayleigh mixture, or the squared modulus $M_j^2$ as a two-state Gamma mixture. This involves estimating the parameters of the distribution functions. For the two-state zero-mean Gaussian mixture model, they are $w_0, \sigma_0$, and $\sigma_1$. For the two-state Gamma mixture model, they are $w_0, \lambda_0$, and $\lambda_1$.

After these parameters are estimated, $p_{S|x}(s = 1|x)$, a posterior probability that a wavelet coefficient is likely to be significant can be obtained following Bayes’ rule. Thus depending on this probability a wavelet coefficients can either approximately keep its original value (if it is significant and attributed to meaningful image features) or shrink towards zero (if it is more likely due to noise). This procedure will be described in the next subsection.

### 6.3.2 Wavelet Coefficients Shrinkage Using Inter-scale Correlation

According to Bayes’ Theorem, given a sample $x$ of the random variable $X$, which follows an $N$-state mixture model defined in Equation (6.73), the posterior probability that $x$ falls in state $s = i$ is determined by

$$p_{S|X}(s = i|x) = \frac{P(s = i)p_{X|S}(x|s = i)}{p_X(x)} = \frac{P(s = i)p_{X|S}(x|s = i)}{\sum_{j=0}^{N-1} P(s = j') p_{X|S}(x|s = j')}$$

(6.80)

where the random variable $X$ can either be the component wavelet coefficients $d_j^H, d_j^V$, or their modulus $M_j$ or squared modulus $M_j^2$. In particular, for the purpose of denoising we are interested in the posterior probability that a wavelet coefficient is significant, namely $p_{S|X}(s = 1|x)$. This probability function can potentially serve as a continuous shrinking function or a “mask” to be imposed on the wavelet coefficients such that the ones due to prominent image features (with high posterior probability $p_{S|X}(s = 1|x)$) are left relatively intact, while the ones due to noise (with low posterior probability) are shrunk towards zero.

Thus at each scale, we can process the component wavelet coefficients image $d_j^H$ using the corresponding posterior probability function $p_{S|d_j^H}(s = 1|x)$ (or process $d_j^V$ using $p_{S|d_j^V}(s = 1|x)$). Another possibility is to use that of the modulus $p_{S|M_j}(s = 1|x)$ or squared modulus $p_{S|M_j^2}(s = 1|x)$ as an overall mask to process both $d_j^H$ and $d_j^V$ at that
scale. In the following text, we use $p_{s|M^2_j}(s = 1|x)$, which is given by

$$p_{s|M^2_j}(s = 1|x) = \frac{P(s = 1)p_{M^2_j|s}(x|s = 1)}{P(s = 0)p_{M^2_j|s}(x|s = 0) + P(s = 1)p_{M^2_j|s}(x|s = 1)}$$

$$= \frac{w_1 \frac{1}{\lambda_1} e^{-x/\lambda_1}}{w_0 \frac{1}{\lambda_0} e^{-x/\lambda_0} + w_1 \frac{1}{\lambda_1} e^{-x/\lambda_1}}$$

(6.81)

Figure 6.13: Illustration of a posterior probability function $m_j$ derived from $M^2_j$. White pixels correspond to high probability close to 1, while black pixels close to 0. (a)-(d) $m_j, j = 1, 2, 3, 4$. 
6.3. **NOISE REDUCTION SCHEME IN THE WAVELET TRANSFORM DOMAIN**

and after rearrangement and notation simplification becomes

$$m_j(x) = \frac{1}{1 + \frac{\lambda_0 \lambda_1}{\lambda_1 \lambda_0} e^{-(\lambda_1 - \lambda_0) x / \lambda_0 \lambda_1}}. \quad (6.82)$$

Figure 6.13 illustrates a posterior probability function $m_j$ derived from $M_j^2$ at four scales (the corresponding $M_j$ are shown in Figure 6.8). White pixels indicate high probability close to 1, while black pixels are close to 0. As we can see, prominent object boundaries are consistently present as clustered white pixels at each scale, indicating the high probability that they truly form parts of object boundaries. However, there are also spurious bright pixels due to noise at lower scales.

To differentiate the legitimate white pixels from the spurious ones, we need to exploit the inter-scale correlation of the wavelet coefficients, and hence the significant values that remain consistent across different scales in a posterior probability functions $m_j$. To exploit this inter-scale correlation, a “mask” function was constructed in [160] using the multi-scale products of the wavelet coefficients of adjacent higher scales and applied to the wavelet coefficients to remove those due to noise.

Intuitively, a “mask” function should be defined on the basis of the probability functions $m_j$ such that for each pixel the mask is close to 1 only if all of the corresponding values in $m_j$ are close to 1 (indicating consistent high probability that it is a legitimate edge point), but is approximately 0 if any of them is close to 0 (indicating a false response due to noise). Following [163], we take the harmonic mean of the probability functions $m_j$, which meets this requirement, and apply it to the wavelet coefficients as a “mask” or shrinkage function. In other words, the inter-scale correlation of the wavelet coefficients are exploited in the form of the harmonic mean of the posterior probability functions $m_j$.

Thus at each scale $2^j$, the “mask” function for shrinking the wavelet coefficients is defined as the harmonic mean of the posterior probability functions $m_j$ from the current scale $2^j$ up to the highest decomposition scale $2^J$, namely,

$$\hat{m}_j(x) = \frac{J - j + 1}{\sum_{j' = j}^{J} \frac{1}{m_{j'}(x)}}. \quad (6.83)$$

Since the probability functions $m_j$ at higher scales are relatively “noise-free”, by taking the harmonic mean from the current scale $2^j$ up to scale $2^J$, this “noise-free” information contained in $m_j$ at higher scales is propagated to the current (low) scale. In this way those wavelet coefficients with consistent high posterior probabilities (the “significant” ones) will have a mask or shrinking factor close to 1, and are thus most likely to keep their original values. In contrast, the posterior probability of those wavelet coefficients due to noise do not match one another at different scales, thus resulting in a shrinking factor approximately 0.

Figure 6.14 shows the wavelet coefficients shrinkage function $\hat{m}_j(x)$ at different scales. As we can see, compared to Figure 6.13, most of the spurious responses from noise have
Figure 6.14: Illustration of wavelet coefficient shrinkage function $\hat{m}_j$ derived as a harmonic mean of posterior probability functions $m_j$ (shown in Figure 6.13). White pixels are close to 1, while black pixels close to 0. (a)-(d) $\hat{m}_j$ at scales $2^j$, $j = 1, 2, 3, 4$.

been greatly suppressed. Next we apply these “masks” to the wavelet coefficients component $d_{jH}$ and $d_{jV}$ at each scale in a top-down manner (that is, higher scales processed first)

$$
\begin{align*}
\hat{d}_{jH} &= d_{jH} \hat{m}_j \\
\hat{d}_{jV} &= d_{jV} \hat{m}_j
\end{align*}
$$

(6.84)

where at the highest scale $\hat{m}_{J} = m_J$ and at other scales $\hat{m}_j$ are derived according to
Equation (6.83) for $1 \leq j \leq J - 1$. Consequently we now obtain the corresponding noise-reduced versions of the wavelet coefficients component, $\hat{d}_j^H$ and $\hat{d}_j^V$. The last step that remains is just reconstruct the original image from $\hat{d}_j^H$, $\hat{d}_j^V$ and the coarse scale information $a_J$, which has been left intact in our denoising procedure.

### 6.3.3 Summary of the Noise Reduction Scheme

Now it is time to summarise our noise reduction scheme in the wavelet transform domain, which includes the following steps:

1. First an image is decomposed via Mallat and Zhong’s dyadic wavelet transform into its “details”, namely, the wavelet coefficients component $d_j^H$ (Equation (6.65)), $d_j^V$ (Equation (6.66)), for $1 \leq j \leq J$, and the “approximations” or scaling coefficients $a_J$ (Equation (6.68)), where $J$ is the highest decomposition scale.

2. Compute the squared modulus $M_j^2$ from $d_j^H$ and $d_j^V$ at each scale according to Equation (6.76).

3. At each scale, model $M_j^2$ as a two-state Gamma mixture as in Equation (6.78) and estimate the hyperparameters $w_0$, $\lambda_0$, and $\lambda_1$.

4. Compute the posterior probability function $m_j$ as in Equation (6.82) using the parameters $w_0$, $\lambda_0$, and $\lambda_1$ estimated in the previous step.

5. Calculate the shrinking function $\hat{m}_j$ as in Equation (6.83). For instance, if $J = 5$, then for scale $2^3$, $\hat{m}_3 = \frac{3}{\frac{3}{m_3} + \frac{1}{m_4} + \frac{1}{m_5}}$.

6. Apply the shrinking function $\hat{m}_j$ to $d_j^H$ and $d_j^V$ according to Equation (6.84) and obtain their filtered versions $\hat{d}_j^H$ and $\hat{d}_j^V$.

7. Reconstruct a noise-reduced image from $\hat{d}_j^H$, $\hat{d}_j^V$, and $a_J$ using Equation (6.69).

### 6.4 Results and Validation

We applied our noise reduction scheme to simulated B-scan ultrasound images. Each image is decomposed up to scale $2^5$. One of the original simulated images has been shown in Figure 6.6(a). Its denoised version (by the noise reduction scheme) is shown in Figure 6.15(a). As we can see, the speckle present in the original image is substantially reduced. In particular, compared to the original speckled image (Figure 6.6(a)), the homogeneous background has been restored very well. So have the foreground circular objects, making the edges between the objects and the background sharp and clear.

Figure 6.15(b) shows the difference between the two images, obtained by subtracting the denoised image from the original one. As we can see, the granular speckles are prevalent in the difference image, the appearance of which is comparable to what we would expect.
SPECKLE REDUCTION USING THE DYADIC WAVELET TRANSFORM

Figure 6.15: The noise-reduced version of the simulated B-scan ultrasound image shown in Figure 6.6(a) and their difference. (a) Noise-reduced after being processed by our noise reduction scheme in the wavelet transform domain. (b) The difference between the original image and noise-reduced image.

Figure 6.16: Comparison of the histogram of the original image and that of the noise-reduced image. (a) Histogram of the original image. (b) Histogram of the noise-reduced image.
6.4. RESULTS AND VALIDATION

(a)

(b)

Figure 6.17: Simulated image (Figure 6.6(a)) denoised by median filters. (a) Denoised by a $5 \times 5$ median filter. (b) By a $9 \times 9$ median filter.

a simulated image of a homogeneous area to look like. This reassures us that most of the speckle in the original image have been removed.

Figure 6.16 compares the histograms of the original image and the noise-reduced image. Note the unimodal appearance of the histogram of the original image (Figure 6.16(a)), which would surely land any histogram-based segmentation methods in trouble. In contrast, Figure 6.16(b) clearly demonstrates bimodal property of the histogram of the denoised image. Obviously the two peaks respectively correspond to the background and the circular objects.

6.4.1 Qualitative Evaluation

We first evaluate our noise reduction scheme qualitatively by comparing the denoised image with those obtained by applying two-dimensional median filters, Wiener filters and Lee filters. Figure 6.17 shows the images filtered respectively by $5 \times 5$ and $9 \times 9$ median filters. As we can see, when applied to a small neighbourhood, e.g. $5 \times 5$ (Figure 6.17(a)), the median filter is not effective in removing the speckle noise present in the simulated image. For a large $9 \times 9$ neighbourhood (Figure 6.17(b)), although it successfully removes speckle noise it also blurs edges present within the image, and leaves such artifacts as streaks or blotches in the background.

The results of applying $5 \times 5$ and $9 \times 9$ Wiener filters are shown in Figure 6.18. The performance of the Wiener filters is no better than that of the median filters. Although the image processed by the $5 \times 5$ Wiener filter (Figure 6.18(a)) looks darker, it is comparable to that processed by a counterpart median filter. The output of the $9 \times 9$ Wiener filter (Figure 6.18(b)) has a cleaner background compared to that processed by the $9 \times 9$ median filter.
Figure 6.18: Simulated image (Figure 6.6(a)) denoised by Wiener filters. (a) Denoised by a $5 \times 5$ Wiener filter. (b) By a $9 \times 9$ Wiener filter.

Figure 6.19: Simulated image (Figure 6.6(a)) denoised by Lee filters. (a) Denoised by a $5 \times 5$ Lee filter. (b) By a $9 \times 9$ Lee filter.
6.4. RESULTS AND VALIDATION

filter. However, the foreground objects are still quite noisy, indicating the speckle has not been reduced effectively in these areas.

Figure 6.19 demonstrates the performance of $5 \times 5$ and $9 \times 9$ Lee filters. Likewise, the result of applying the $5 \times 5$ Lee filter (Figure 6.19(a)) is not satisfactory at all. The performance of the $9 \times 9$ Lee filter is better, but it still has difficulties in dealing with the speckle around high contrast areas such as object boundaries, as evidenced by the “residue” speckle left over.

In contrast, our noise reduction scheme deals very well with both the background and the foreground objects, leaving a nearly “pure” background and superbly restored foreground objects. Thus it is not unfair to say that it is superior to either median, Wiener or Lee filters in removing the speckle and giving a visually better denoised image.

6.4.2 Quantitative Validation

Except for the commonly used signal-to-noise ratio (SNR), there is no other definitive or widely accepted quantitative evaluation criteria available for assessing the success of such noise reduction procedures. Following [21, 24], the SNR of an image is defined as the ratio of the mean and standard deviation of the image pixels (see section 3.1). Apparently, the more homogeneous an image, the larger the corresponding SNR. In other words, we would expect an improvement in SNR from any noise removal methods applied.

We next validate our noise reduction scheme quantitatively using SNR. Table 6.2 compares the SNR of simulated ultrasound images and the corresponding denoised images respectively by our noise reduction scheme (NRS), $5 \times 5$ median filter (MED5), $9 \times 9$ median filter (MED9), $5 \times 5$ Wiener filter (WEI5), $9 \times 9$ Wiener filter (WEI9), $5 \times 5$ Lee filter (LEE5), and $9 \times 9$ Lee filter (LEE9). The first five rows of the table list the SNR for five randomly selected images (denoted as 1-5). The last two rows are the mean (denoted as MEAN$_{100}$) and standard deviation (STD$_{100}$) of the SNR of 100 simulated images.

As we can see from Table 6.2, our noise reduction scheme is second to none among the methods in the comparison. The SNR is increased to 3.914 on average, which is a substantial improvement over that of the original, 1.886. The $9 \times 9$ Lee filter is the second best, 3.523 on average, followed by the $9 \times 9$ median filter, with $5 \times 5$ Wiener filter being the worst.

Generally speaking, for median, Wiener and Lee filters, as the window size increases, so does the SNR of the denoised images. However, this relation between filter size and improvement in SNR is not always true. There is a point beyond which no more improvement is possible by increasing filter size. For example, we have found the performance of a $19 \times 19$ Lee filter is no better than its $9 \times 9$ counterpart. A similar situation holds true for median and Wiener filters. There is no way for these filters to surpass our noise reduction scheme by increasing filter size.

SNR is equivalent to what is known as ENL (equivalent number of looks) in the SAR imaging community.
SPECKLE REDUCTION USING THE DYADIC WAVELET TRANSFORM

<table>
<thead>
<tr>
<th>IMAGE</th>
<th>ORI</th>
<th>NRS</th>
<th>MED5</th>
<th>MED9</th>
<th>WIE5</th>
<th>WIE9</th>
<th>LEE5</th>
<th>LEE9</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1.823</td>
<td>3.866</td>
<td>3.031</td>
<td>3.290</td>
<td>2.893</td>
<td>3.102</td>
<td>2.945</td>
<td>3.377</td>
</tr>
<tr>
<td>4</td>
<td>1.892</td>
<td>3.961</td>
<td>3.101</td>
<td>3.373</td>
<td>2.957</td>
<td>3.176</td>
<td>2.921</td>
<td>3.600</td>
</tr>
<tr>
<td>5</td>
<td>1.887</td>
<td>3.911</td>
<td>3.039</td>
<td>3.311</td>
<td>2.935</td>
<td>3.156</td>
<td>3.042</td>
<td>3.540</td>
</tr>
<tr>
<td>MEAN100</td>
<td>1.886</td>
<td>3.914</td>
<td>3.048</td>
<td>3.312</td>
<td>2.928</td>
<td>3.145</td>
<td>3.031</td>
<td>3.523</td>
</tr>
<tr>
<td>STD100</td>
<td>0.008</td>
<td>0.052</td>
<td>0.032</td>
<td>0.039</td>
<td>0.027</td>
<td>0.034</td>
<td>0.082</td>
<td>0.077</td>
</tr>
</tbody>
</table>

Table 6.2: Comparison of the signal-to-noise ratio (SNR) of the original simulated B-scan ultrasound images (ORI) and the corresponding denoised images by our noise reduction scheme (NRS), 5×5 median filter (MED5), 9×9 median filter (MED9), 5×5 Wiener filter (WIE5), 9×9 Wiener filter (WIE9), 5×5 Lee filter (LEE5), and 9×9 Lee filter (LEE9).

6.4.3 Application to Real Ultrasound Images

We finally apply our noise reduction scheme to real ultrasound images of the prostate and one of the results is shown in Figure 6.20. The original ultrasound image of the prostate “us1”, which is degraded by speckle noise, is shown in Figure 6.20(a). Figure 6.20(b) shows its denoised version. As we can see, most of the speckle noise is significantly suppressed, especially in homogeneous areas, while the important features of the original image are preserved. Note the two small bright regions within the prostate are enhanced after noise reduction.

The speckle removed is evident in Figure 6.21, which shows the difference between the original ultrasound image “us1” and that denoised by the noise reduction scheme. As we can see, the granular appearance dominates most parts of the image except the signal-free regions at the four corners. Moreover, the white markers along the boundaries of the image have also been suppressed and end up in the difference image. This is understandable as their sizes are almost the same as that of the granular speckle and the noise reduction scheme cannot discriminate them.

Figure 6.22 shows another example of application of the noise reduction scheme. The difference between the original ultrasound image “us5” and that denoised by the noise reduction scheme is shown in Figure 6.23. Table 6.3 compares the SNR of some of the ultrasound images of the prostate and the corresponding denoised images respectively by our noise reduction scheme (NRS), 5×5 median filter (MED5), 9×9 median filter (MED9), 5×5 Wiener filter (WIE5), 9×9 Wiener filter (WIE9), 5×5 Lee filter (LEE5), and 9×9 Lee filter (LEE9). As we can see from our noise reduction scheme surpasses all these methods in the comparison.
6.5 Conclusion

Speckle is an undesired artifact that is inherent to optical, synthetic aperture radar, and ultrasound imaging. Many adaptive filtering methods based on local image statistics have been proposed for speckle reduction in the last few decades and have their own limitations (refer to Chapter 5 for details). The advent of the wavelet framework has opened up a new frontier in this area. Of special interest to us is Mallat and Zhong’s dyadic wavelet transform (MZDWT), which features a quadratic spline wavelet and scaling function.

MZDWT is similar to a conventional discrete wavelet transform involving biorthogonal wavelets but without downsampling. It provides an overcomplete representation of a
signal in the wavelet transform domain. It is this information redundancy that makes it translation-invariant, one of the most noteworthy properties of this transform. In contrast, the conventional discrete wavelet transforms are translation-variant due to downsampling. For these transforms extra effort such as cycle spinning is needed to achieve comparable translation invariance [98, 230], whereas for MZDWT it is a “built-in” feature.

Another feature of this transform is that the quadratic spline wavelets and the associated band-pass filters are a multi-scale “edge detector”. The resultant component wavelet coefficients constitute gradient vectors, the modulus of which forms multi-scale edge maps of an image. Important image features such as object boundaries are consistently present as local extrema, which are aligned at the same positions across different scales due to translation invariance. On the contrary, the response due to noise are inconsistent and quickly subside as scale increases. Thus this inter-scale correlation (or uncorrelation for noise) can be exploited to remove noise.

Inspired by [163], we proposed in this chapter a new speckle noise reduction scheme in the wavelet transform domain, which is on the basis of modelling the statistical properties of the wavelet coefficients, and takes advantage of their inter-scale correlation. Especially, we modelled the squared modulus of the component wavelet coefficients as a two-state Gamma mixture. Inter-scale correlation is exploited by taking the harmonic mean of the posterior probability functions, which are derived from the Gamma mixture.

We applied this noise reduction scheme to both simulated and real ultrasound images, and its performance is quite satisfactory. The important features of the original noise corrupted image are preserved while most of the speckle noise is removed successfully. We evaluated it both qualitatively and quantitatively by comparing it with median, Wiener, and Lee filters. The results showed that it definitely surpasses all these filters in comparison.
6.5. CONCLUSION

Figure 6.22: Another example of application of the noise reduction scheme. (a) The original ultrasound image “us5”. (b) After denoised by the noise reduction scheme.
Figure 6.23: Difference between the original ultrasound image “us5” (Figure 6.22(a)) and that denoised by the noise reduction scheme (Figure 6.22(b)).

<table>
<thead>
<tr>
<th>IMAGE</th>
<th>ORI</th>
<th>NRS</th>
<th>MED5</th>
<th>MED9</th>
<th>WIE5</th>
<th>WIE9</th>
<th>LEE5</th>
<th>LEE9</th>
</tr>
</thead>
<tbody>
<tr>
<td>us1</td>
<td>2.39</td>
<td>2.82</td>
<td>2.54</td>
<td>2.60</td>
<td>2.55</td>
<td>2.65</td>
<td>2.47</td>
<td>2.58</td>
</tr>
<tr>
<td>us4</td>
<td>2.16</td>
<td>2.43</td>
<td>2.25</td>
<td>2.29</td>
<td>2.27</td>
<td>2.32</td>
<td>2.23</td>
<td>2.27</td>
</tr>
<tr>
<td>us5</td>
<td>2.13</td>
<td>2.35</td>
<td>2.19</td>
<td>2.22</td>
<td>2.22</td>
<td>2.26</td>
<td>2.12</td>
<td>2.20</td>
</tr>
<tr>
<td>us6a</td>
<td>2.31</td>
<td>2.67</td>
<td>2.51</td>
<td>2.56</td>
<td>2.52</td>
<td>2.66</td>
<td>2.44</td>
<td>2.53</td>
</tr>
<tr>
<td>us6b</td>
<td>2.74</td>
<td>3.92</td>
<td>3.19</td>
<td>3.37</td>
<td>3.11</td>
<td>3.37</td>
<td>2.87</td>
<td>2.95</td>
</tr>
</tbody>
</table>

Table 6.3: Comparison of the signal-to-noise ratio (SNR) of the original ultrasound images of the prostate (ORI) and the corresponding denoised images by our noise reduction scheme (NRS), 5 × 5 median filter (MED5), 9 × 9 median filter (MED9), 5 × 5 Wiener filter (WEI5), 9 × 9 Wiener filter (WEI9), 5 × 5 Lee filter (LEE5), and 9 × 9 Lee filter (LEE9).
6.5. CONCLUSION
Chapter 7

Contour Detection Using the Laplacian of Gaussian Operator

In computer analysis of prostate ultrasound images, detection of the contour of the prostate is difficult because of the ultrasound images’ low resolution and high level of noise. In this chapter we present a method for detecting the contour of the prostate by employing the Laplacian of Gaussian (LoG) or Marr-Hildreth operator, the Gaussian kernel of which acts as a low pass filter eliminating the high frequency noise. Convolution of the operator with the ultrasound images is carried out by use of fast Fourier transform (FFT) to reduce the computational overhead. LoG operators of various sizes have been applied to the ultrasound images and the results are encouraging.

7.1 The Laplacian of Gaussian (LoG) Operator

Many segmentation methods depend on reliable edge detection at an early stage. Edge points are pixels at locations where abrupt local gray scale changes or discontinuities occur, such as the boundary between two regions of relatively distinct average intensity or texture. The development and implementation of edge detection schemes have been a major pursuit of the computer vision and image processing community [223, 231–239].

The well established edge detectors, such as the Sobel, Prewitt, and Roberts operators, essentially involve a differentiation process [117, 240, 241]. Depending on the order of derivative taken, these edge detectors can be classified into two main categories: gradient operators and second derivative operators. For the gradient operators, edges are detected as local maxima of gradients of intensity changes. For the second derivative operators, edges are detected as zero-crossings of the second derivative. The Laplacian operator is a frequently used rotationally symmetric or isotropic second order operator. Unfortunately, these operators are very sensitive to high frequency noise because of the differentiation process involved [240]. Thus they are not very useful for images of low signal-to-noise ratio (SNR) without pre-processing such as low-pass filtering to reduce the noise.

Combining a low pass filter and a Laplacian operator, the Laplacian of Gaussian (LoG)
operator or Marr-Hildreth operator was first proposed by Marr and Hildreth [212,242,243] in the 1980s based on their investigation of the human visual system. A Gaussian is employed to smooth an image before the Laplacian is applied. It is believed that the LoG operator closely matches the first stage neural filter of biological vision systems [243,244].

Formally, an image \( I(x, y) \) is first smoothed by a Gaussian,

\[
G_\sigma(x, y) * I(x, y)
\]  

(7.1)
where $\ast$ denotes convolution, and $G_{\sigma}(x, y)$ is a two dimensional Gaussian filter

$$G_{\sigma}(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

(7.2)

with Fourier transform

$$g_{\sigma}(u, v) = e^{-\frac{1}{2}\sigma^2(u^2+v^2)}$$

(7.3)

where $\sigma$ is the standard deviation of the Gaussian filter, which defines the extent over which smoothing is carried out.

A Laplacian operation is then applied to the result of the above convolution. Since

$$\nabla^2[G_{\sigma}(x, y) \ast I(x, y)] = \nabla^2G_{\sigma}(x, y) \ast I(x, y)$$

(7.4)

the two operations can be combined into one filter, namely, the LoG operator, which is given by

$$\nabla^2G_{\sigma}(x, y) = \frac{1}{\pi\sigma^4} \left(1 - \frac{x^2+y^2}{2\sigma^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}.$$  

(7.5)

Considering Equations (7.2), (7.3), and

$$\mathcal{F}\{f(x, y)\} \Leftrightarrow F(u, v)$$

$$\mathcal{F}\{\nabla^2f(x, y)\} = -2\pi^2(u^2+v^2)F(u, v)$$

(7.6)

where $\mathcal{F}$ denotes the Fourier transform and $f(x, y)$ and $F(u, v)$ is a Fourier transform pair, we obtain the Fourier transform of the LoG operator

$$\mathcal{F}\{\nabla^2G_{\sigma}\}(u, v) = -2\pi^2(u^2+v^2)e^{-\frac{1}{2}\sigma^2(u^2+v^2)}.$$  

(7.7)

Figure 7.1 illustrates a LoG operator with $\sigma = 1$. Its two-dimensional graph is shown in Figure 7.1(a), along with a one-dimensional profile through its centre in Figure 7.1(b). The size of a LoG operator is determined by $w$, the diameter of the central positive region, $w = 2\sqrt{2}\sigma$, as illustrated in Figure 7.1(b). Although it is suggested that the LoG operator should be implemented over a minimal support of $4w$ to reduce distortion introduced by truncation of the operator in discrete form [242, 245], we found a support of $3w$ sufficient for our application. Refer to [245–249] for computational techniques and implementation considerations for the LoG operator.

By convolving an image with a LoG operator, edges can be located as the loci of points where the second directional derivative crosses zero in the output of the convolution. In this way, edges occurring at different resolutions can be obtained by applying LoG operators with different standard deviations, because the corresponding embedded Gaussian filter selects the range of scales over which intensity changes take place [242].
7.2 Contour Detection Based on the LoG Operator

As a simple implementation, an ultrasound image is first convolved with a LoG operator, the result of which is then binary thresholded, namely, the positive values are coloured white and the negative values black. Thus the zero-crossings can then be located as the boundary pixels between the black and white regions (say by applying Prewitt edge detectors). Figure 7.2 illustrates the binary thresholded images after LoG operators of varying $\sigma$ are applied to image “us1” (shown in Figure 2.9(b)).

As we can see, the contour of the prostate is “emerging” as $\sigma$ increases, which is quite encouraging. However, there are several issues that need to be addressed. Firstly, due to the low resolution and high level of noise of the ultrasound images of the prostate it is necessary to use LoG operators with $\sigma$ larger than 10, typically with spans of more than $85 \times 85$ elements, in contrast to $3 \times 3$ masks of traditional edge detectors. This makes a brute force convolution too computationally expensive.

Another problem is the severe border erosion effect due to the convolution process, as demonstrated in Figure 7.2. Generally, when an image is convolved with an operator of size $M \times M$ elements, $\lceil M/2 \rceil$ pixels from either side of the image borders will be lost,
CONTOUR DETECTION USING THE LAPLACIAN OF GAUSSIAN OPERATOR

where \([x]\) denotes the largest integer less than or equal to \(x\).

Zero-padding before the convolution is one of the usually used method for mitigating the border erosion effect, however, this will produce false intensity changes, especially false step edges if there are bright regions at the border of original image, interfering with the true zero crossings there. Instead, for the purpose of contour detection, before convolution an image should be extended with boundary pixels being copied outward \([M/2]\) times, which has been shown to be superior to zero padding [245].

To address the first problem we resort to the convolution theorem, which suggests that large convolution calculations can often be performed more efficiently by multiplication in the frequency domain [241]. The convolution in Equation (7.4) can then be accomplished using the fast Fourier transform (FFT). Because the Fourier transform of the LoG operator is analytically available, as given in Equation (7.7), and thus can be computed beforehand, there are only two FFTs involved for computation of the convolution, considerably reducing the computation cost.

Another problem with the above implementation is that binary thresholding simplifies the task of locating zero-crossings, but the information on the strength or magnitude of candidate edge points has been ignored. Therefore some edges of little importance and even spurious edges are introduced, especially for the cases of low \(\sigma\) value, as we can see in Figure 7.2.

Associated with each zero crossing is its slope, which is related to the contrast and span of the local intensity change, namely, the magnitude of the edge. The edge detection scheme above can then be improved by incorporating an edge strength detector, which measures the magnitude of slope for each zero crossing, into the procedure of locating the edge points by use of the LoG operator. Thus a threshold can be set to remove those “weak” or spurious edge points whose strength are below it. The strength of each candidate edge point can be obtained by computing the maximum slope in the four possible directions of its \(3 \times 3\) neighbourhood.

In summary, to detect the contour of the prostate, an ultrasound image is first padded with boundary pixels, and then convolved with a LoG operator using the FFT. From the output of the convolution edge points are found as zero crossings and their corresponding strengths are computed and thresholded to give the final edge map.

7.3 Results and Discussion

The result of applying the LoG operator based edge detection scheme to ultrasound image “us1” is shown in Figure 7.3. Figure 7.3(a) is the original image (also shown in Figure 2.9(b)) and Figure 7.3(b) is the result after applying the LoG operator (\(\sigma = 12.61\)), with the detected edges superimposed on the original image. As we can see, the detected contours closely match the boundaries of the regions discernible by the eye, including not only most parts of the prostate itself, but also some features inside and outside it. We can
7.3. RESULTS AND DISCUSSION

Figure 7.3: Result of applying the LoG operator to the image “us1”. $\sigma=12.61$. (a) Original image “us1”. (b) Contour detected. (c) Weak edge points removed.
also see that all the edge segments are presented as closed contours. There are also some
spurious edges, such as those little irregular circles, which are difficult to justify.

Figure 7.3(c) shows the result of the improved implementation, with the edge strength
detector incorporated. After thresholding at 0.85 (normalised value), many weak or spu-
rious edge points have been removed. Figure 7.4 shows another example.

As we have seen, the Gaussian kernel of the LoG operator serves as a low-pass filter.
It is argued that the Gaussian is an optimal smoothing filter since it optimally satisfies
two contradictory constraints which require that the span of the filter be as narrow as
possible in both the spatial domain and the frequency domain [232, 242]. The larger the
LoG operator, that is, the larger the $\sigma$, the more noise will be removed, but at the same
time, the less detail will be preserved. A compromise must be made between eliminating
noise and preserving local details. Generally the contours of features smaller than $w$
are displaced and may become fused with the contours of nearby features [250].

Moreover, the LoG operator tends to produce closed contours, which is a desirable
feature in some cases where global edges are dominant and to be pursued. Unfortunately,
if there are strong noise blobs or other prominent small features, they can easily interfere
with the continuity of the sought-after contours, displacing them to different closed curves.

For instance, as shown in Figure 7.3, the hyperechoic region (the small white region
within the left part of the prostate) produces an axial shadow, which consequently gen-
erates a gap and interrupts the prostate contour at the upper left side. Without this
region, the contour there would be continuous. For these cases, closed contours produced
by the LoG operator are hardly an advantage since the component edge segments of the
prostate contour might lie on different closed curves. This makes it very difficult to trace,
interpolate, and link those legitimate edge segments to yield a meaningful outline of the
prostate, which can be done only if some a priori knowledge about the prostate and the
images is incorporated into the process, such as prostate anatomy and its general position,
shape and relative size within the images.

7.4 Conclusion

In summary, as a pilot study, we explore in this chapter the application of the LoG
operator to detecting the contour of the prostate. Although the results are encouraging,
the LoG operator alone is inadequate to produce a meaningful closed contour of the
prostate. Nevertheless, as we have seen in Figure 7.2, LoG operators of various $\sigma$ generate
contours in a fine to coarse manner, which is a valuable experience lending us a multi-scale
perspective, as we will detail in Chapter 10. Before that, let us review in Chapter 8 the
current literature on the problem of extracting the prostate contour.
Figure 7.4: Result of applying the LoG operator to the image “vs04”. $\sigma=11.67$. (a) Original image “vs04”. (b) Contour detected. (c) Weak edge points removed.
Chapter 8

Delineation of Prostate Contours: A Critical Review

It is essential to accurately extract prostate contours in many diagnostic and treatment procedures for prostate disease. However, due to low resolution, poor contrast and high noise levels, it is a difficult task to locate the contour of the prostate in ultrasound images automatically by computer. Only a handful of limited successes have been reported. This chapter reviews the literature on the problem of delineating prostate contours in B-scan ultrasound images.

8.1 Introduction

Accurate delineation of prostate contours is crucial for many diagnostic and treatment procedures based on TRUS imaging. For example, serum prostate-specific antigen density (PSAD\(^1\)) is very useful in differentiating benign from malignant prostate disease [251, 252]. Determination of PSAD involves the measurement of the prostate gland volume. Inaccurate volume estimates may potentially put a big question mark over the value of PSAD in early detection of prostate cancer.

Three methods have been proposed for prostate volume determination using TRUS imaging. Among them planimetric volumetry [253–255] provides the most accurate and reproducible correlation with the actual gland size. In this approach, the prostate is scanned from apex to base (or in the opposite direction) by moving the ultrasound probe at a sub-centimeter interval (commonly from 0.25cm to 0.5cm) and transverse images are obtained for each step. The boundaries of the prostate are manually outlined in these sequential cross-sectional images. The areas enclosed by the boundaries are then calculated and the volume of each section is obtained. The total volume can be estimated by summing up the volumes of each section. Unfortunately manual tracing of the prostate

\(^1\)Prostate-specific antigen density is defined as prostate-specific antigen (PSA) in proportion to clinical prostate gland volume (CV). \(PSAD = \frac{PSA}{CV}\).
contour is a time consuming and labor intensive procedure, and so planimetric volumetry has not found wide application in clinical practice.

In the second method the prostate volume is estimated by assuming the prostate gland to have the shape of a sphere, spheroid, or ellipsoid, and applying the corresponding volume calculation formula for these shapes [254]. Gland dimensions such as latero-lateral, antero-posterior and cranio-caudal dimension need to be measured. Although accuracy of this method is inferior to planimetric volumetry, it is the most widely used and accessible method in clinical practice because of its fast and simple application. The third method relies on the determination of the maximum transverse area of the prostate, which is manually outlined, and of the cranio-caudal dimension of the gland [255].

Both of the latter two methods are nonplanimetric. Bazinet et al. [256] assessed the accuracy and reproducibility of nonplanimetric TRUS volume estimates and concluded that they should be interpreted with caution when used for PSAD calculations in the early detection of prostate cancer because up to 25% volume differences can be expected between two such estimates, despite the fact that consecutive gland volume estimates show statistically good correlation.

Thus accurate delineation of the prostate contours is essential to two of the three prostate volume estimation methods. Contour information of the prostate is also helpful in directing biopsy needles to the suspicious region during diagnosis [6], guiding radioactive seed implantation of the prostate during brachytherapy [257], and monitoring volume change of hypo-echoic lesions to predict and detect disease progression following radiation therapy [258]. Identifying the contour of the prostate in a sequence of ultrasound images may also help restrict the region of interest for tissue characterisation or form a three-dimensional visualisation of the prostate.

Currently the boundaries of the prostate are manually outlined by a urologist, which is a tedious and challenging task [254]. Moreover, manual delineation is subjective and hardly reproducible. Because of the variability in the manual tracing of the prostate contour, repeated volume determination of the prostate by the same urologist often leads to large variations. The results of manual tracing vary even more when done by different urologists [259, 260].

The problems with the manual delineation suggest that automatic contour extraction (or semi-automatic with minimal interaction from urologists) by computer would be of great value. Computer analysis may provide a more reproducible interpretation of ultrasound images of the prostate and locate textual information that human observers may find difficult to perceive because of the limited number of gray scales human eyes can differentiate. It may also be able to considerably reduce the bias a human operator implicitly bears in manual tracing. Should the contour delineation process be automated, wide application of planimetric volumetry would be expected because of its advantage over the other two methods in terms of accuracy.

Unfortunately, the task of determining the prostate boundaries automatically by com-
puter usually is complicated by the low resolution and poor contrast of the B-scan ultrasound images, accompanied by the high level of speckle noise due to scattering and other complicated interactions between ultrasonic pulses and human tissue. Moreover, missing boundary segments are not uncommon, which are usually attributed to acoustic shadowing, and/or hyperechoic structures within or around the prostate gland.

Due to these difficulties only a handful of limited successes have been reported in the literature on automatic or semi-automatic prostate contour extraction by computer. Based on the underpinning strategies, they can be classified into three groups: pixel or texture based [261, 262], edge based [263–267], and model based [268–270]. Table 8.1 summarises the main features and disadvantages of these solutions.
### Table 8.1: Summary of cited references on the problem of prostate contour extraction.

<table>
<thead>
<tr>
<th>STRATEGY</th>
<th>AUTHOR</th>
<th>MAIN FEATURES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
</table>
| Pixel    | Prater [261] | ● A feed-forward neural network used as a pixel classifier  
 ● Images segmented into prostate and non-prostate regions | ● Extensive training needed for the neural network  
 ● Training images need to be segmented manually  
 ● Neural network sensitive to training images |
| Texture  | Richard [262] | ● Four micro-texture masks used to generate four texture energy measures  
 ● Pixels classified into different classes by an unsupervised clustering process according to the energy measures | ● The number of classes produced not predictable  
 ● The classes produced need to be assigned to prostate and non-prostate regions |
| Edge     | Richard [263] | ● Laplacian-of-Gaussian operator carried out in two separate steps  
 ● Initialisation points identified and fitted to establish a guide curve  
 ● An edge selection algorithm to determine the boundary points based on distance to the guide curve | ● User input needed to establish a reference guide, fill gaps and correct any errors of the edge selection algorithm  
 ● Cannot adjust itself adaptively to images of varying contrast |
Table 8.1: Summary of cited references on the problem of prostate contour extraction (continued).

<table>
<thead>
<tr>
<th>STRATEGY</th>
<th>AUTHOR</th>
<th>MAIN FEATURES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
</table>
| Edge     | Aarnink [264]| • Uniform blurring and nonlinear Laplace filtering to generate a zero-crossing image and a gradient image  
• An edge intensity image obtained by masking the gradient image with the zero-crossing image  
• A radial search strategy for the anterior part and a linear search strategy for the posterior part  
• Knowledge based interpolation to fill up the missing parts | • Too many parameters that have to be empirically tuned  
• The nonlinear Laplace filter tends to generate too many closed curves from which the prostate contour is difficult to identify  
• Problematic interpolation and edge linking process; human intervention usually required  
• Contour obtained is not smooth and does not follow the contour discernible to the eye |
|          | Pathak [266,267]| • Edge detection problem modeled as a line segment finding process  
• Contrast enhancement by techniques known as “sticks” and weak membrane fitting  
• Canny’s edge detector applied to generate an edge map  
• Boundary points searched in the radially outward direction from a reference point | • Compromise to be made with respect to stick size selection between speckle reduction and edge enhancement  
• No meaningful closed contour finally produced by the method  
• Manual edge linking and editing totally left to the user |
Table 8.1: Summary of cited references on the problem of prostate contour extraction (continued).

<table>
<thead>
<tr>
<th>STRATEGY</th>
<th>AUTHOR</th>
<th>MAIN FEATURES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Hendrikx [268]</td>
<td>• A number of boundary points specified by the user</td>
<td>• The six template contours not necessarily representative</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Best fit of six predefined template contours found as the prostate contour</td>
<td>• Requires extensive human interaction to designate the boundary points</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Fitting process computationally expensive</td>
</tr>
<tr>
<td></td>
<td>Ladak [269, 270]</td>
<td>• Four end points identified by the user on the prostate boundary</td>
<td>• Problematic initial contour after interpolation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Hermitian interpolation of the four end points leads to an initial crude</td>
<td>• Trapped by local minima of the extraneous edge points corresponding to</td>
</tr>
<tr>
<td></td>
<td></td>
<td>estimate of the prostate shape</td>
<td>surrounding tissues or other artifacts</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Discrete dynamic contour initialised and deforms to fit the prostate contour</td>
<td>• Interactive guidance from the user required to push the snake for a better</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>fit</td>
</tr>
</tbody>
</table>
8.2 Pixel or Texture Based Strategies

The pixel or texture based strategies have an implicit presumption that pixels of the prostate region and non-prostate region belong to different mutually exclusive classes.

8.2.1 Prater’s Feed-Forward Neural Network

Prater and Richard [261] used a feed-forward neural network as a pixel classifier to segment prostate ultrasound images into prostate and non-prostate regions. However, promising results can be expected only after the network is extensively trained using representative teaching image sets, which have to be segmented manually by an expert sonographer. The performance of the neural networks trained on one set of images may be in doubt when applied to a different groups of images.

8.2.2 Richard’s Texture Based Segmentation

A texture-based segmentation method was proposed by Richard and Keen [262]. This method is based on four texture energy measures associated with each pixel in the image, which form the feature space. The four texture energy measures are obtained by convolving the original image with four $5 \times 5$ micro-texture masks, each of which emphasizes a different band of spatial frequencies with different peak gains. An unsupervised iterative clustering process labels each pixel and assigns it to the most probable class according to the energy measures. However, the number of classes produced is not entirely predictable, which poses a new problem to be solved, that is, how to further assign these classes as prostate regions and non-prostate regions.

The segmentation results generated indicate that this texture-based method alone is inadequate for segmenting the ultrasound prostate images. To give more meaningful results it is necessary to combine it with other segmentation methodologies such as edge identification. Being computationally intensive is another disadvantage of this method. On the other hand, the use of the four frequency sieving micro-feature masks has conveyed an implicit positive message that it might be promising to consider the contour extraction problem from a multi-scale point of view, which is also supported by our observations to be described in the next chapter.

8.2.3 Comments on Pixel or Texture Based Segmentation

The difficulties facing the pixel or texture based methods could be attributed to their implicit presumption, that is, pixels of the prostate region and non-prostate region are distinct from each other and belong to two mutually exclusive classes in terms of gray scale values or related statistical characteristics. However, this presumption seldom holds in a real world ultrasound image. This could be explained by the fact that the prostate and surrounding tissues usually have similar acoustic properties. Moreover, the ultrasound
radio frequency (RF) echo signal collected by the probe has undergone many processing steps and system effects before the final gray scale image is formed and presented.

These processing steps include linear or nonlinear processing steps, such as time gain compensation (TGC), which could be overdone, and envelope detection. There also exist system dependent effects such as diffraction due to the finite aperture of the transducer, and tissue-dependent changes in the signal attributed to the frequency dependent attenuation of soft tissue. These effects can only be compensated for, if possible, at the RF stage of the echo signal, before the final gray scale image is formed. Because of these effects, two echo signals which originally contain mutually exclusive information might present themselves as two pixels of the same gray scale in the final image. On the other hand, for the same reason, the gray scale presentation of the same tissue may not be reasonably homogeneous. For this reason several researchers have proposed the use of the unprocessed RF signal or its frequency spectrum in order to find an empirical methodology to characterise prostate tissue [271–274].

8.3 Edge Based Strategies

The edge based strategies [263–267] are based on the observation that the interior of the prostate is hypoechoic compared to the tissues surrounding it, which appear relatively hyperechoic. The boundaries of the prostate can then be identified as the points where the greatest gray level transition occurs.

However, conventional gradient edge detectors, such as the Sobel, Roberts, and Prewitt operators are doomed to failure when applied to the ultrasound prostate images, due to low resolution, poor contrast and the ubiquitous speckle noise, which result in an excessive number of spurious small-scale edge points.

The poor quality of the ultrasound images makes it necessary to reduce speckle and other high frequency noise and improve the contrast before any edge finding algorithm is applied. Chen and his coworkers [265] attempted to smooth the images using direction-based low pass filtering, followed by derivative edge detectors and finally edge linking in which a priori knowledge is employed to obtain prostate contour.

8.3.1 Richard’s Method Based on the Laplacian-of-Gaussian Operator

In their endeavour to form a 3D presentation of the prostate, Richard and his colleagues [263] introduced a method for segmenting ultrasound images based on the Laplacian-of-Gaussian (LoG) operator. Instead of implementing the LoG operator as a single entity, they carried it out in two steps: first as a $41 \times 41$ Gaussian filter convolved with an original ultrasound image, then as a $3 \times 3$ Laplacian operator convolved with the Gaussian filtered image. The resultant image was then thresholded at zero to give a binary image, within which the candidate edge points were identified as the demarcation points between white regions and black regions. The user was required to select several initialisation
points, through which a closed curve was fitted to establish a guide. Using this guide, an edge selection algorithm determined which points were truly edge points lying on the boundary of the prostate, based on distance to the guide curve, direction of sign change and connectivity with other edge points already identified.

This method is far from a complete automatic solution to the segmentation problem by itself, since user input is required to establish a reference curve to remove any extraneous edge segments produced by the LoG operator, to correct any errors arising from the edge selection algorithm, and to fill in gaps, which are not uncommon. Moreover, the authors acknowledged later that this algorithm cannot adjust itself adaptively to cope with the varying contrast from one image to another [262].

8.3.2 Aarnink’s Practical Clinical Method for Contour Determination

Aarnink and his colleagues [264] designed a fast and easy-to-use prostate contour determination method intended for use in everyday clinical volume measurements. This technique is based on nonlinear Laplace filtering where edge points are identified as zero-crossings of the second derivative of the image. A uniform blurring filter is first applied to smooth the image for noise reduction. The filtered image is then sub-sampled by a factor of 3 to accelerate the processing steps that follow, such as differentiation, edge enhancement and selection, and finally edge linking by interpolation.

A nonlinear Laplace filter, that is, the second derivative in gradient direction (SDGD) method, is used as the differentiation operator. It was implemented using local min-max filters to approximate the SDGD. The zero-crossing image and gradient image are generated from the differentiation process. Next the gradient image is masked with the zero-crossing image to give an edge intensity image. Edge points are marked as local maxima in the edge intensity image after thresholding with hysteresis.

A relaxation technique is applied to make sure that weak edge points directly connected to stronger ones are enhanced. These detected edge points will be the starting points for a search in the neighbourhood for other potential edge points that form parts of the prostate boundary. A radial search strategy is used for the anterior (upper) part of the prostate and a linear search strategy for the posterior (lower) part. The search is based on the centre of the prostate, which is manually indicated by a user. Finally the edge points are linked to form a closed contour, after knowledge-based interpolation to fill up the missing parts, and removal of side branches at bifurcations.

Although this method is fast to implement and easy to use, there are several issues that cause concern.

- Firstly, there are too many parameters that have to be empirically adjusted, such as the size of the uniform blur filter, the size of the min-max filter and the predefined value in thresholding with hysteresis. Performance of this method thus greatly depends on the selection of the parameters and their combined effects, and will not be consistent. For example, when choosing the min-max filter size, a compromise has
to be made between two conflicting expectations, that is, noise reduction and edge localisation accuracy. Filters of larger size produce more globally significant edges, and noise interference will be kept to a minimum. The edge points will be more accurately localised when a smaller filter is applied, however, at the expense of noise reduction.

- Secondly, similar to the Laplacian of Gaussian filter, the nonlinear Laplace filter tends to generate closed contours. This property would be an advantage only if ideally the detected boundary forms one closed curve. Unfortunately, this is seldom the case in practice. It is most likely that the prostate boundary will consist of several component segments which are parts of several meaningless closed contours, which have no self-evident correspondence with any anatomical objects. This can be attributed to the uniform blur filtering, which generally displaces features smaller than the filter size, causing small features close to one another to be merged.

- Moreover, the final interpolation and edge linking process remains problematic since it is very difficult to fill in the gaps between boundary segments and remove the false branches introduced during the interpolation process. Human intervention is necessary for some cases where automatic interpolation and pruning fail. Finally, another disadvantage of this method is that the contour obtained is far from locally smooth. It is rather jagged and does not reasonably follow the contour discernible to human eyes, especially for the parts obtained through interpolation.

**8.3.3 Pathak’s Edge-Guided Boundary Delineation**

With the hypothesis that inter-observer variability can be lowered with edge guidance provided by computer, Pathak et al. [266, 267] described a semi-automatic boundary delineation paradigm. Candidate edge points are detected and presented as a visual guide to the operator, who then manually outlines the prostate boundary.

Firstly, the contrast of the image is enhanced and speckle is reduced by applying a special filtering technique using line segments known as “sticks” [275, 276]. Here the edge detection problem is modeled as a line finding process rather than a step change finding process underpinning the conventional gradient edge operators, where edges are located at the points where abrupt gray scale transitions occur. Within a predefined small neighbourhood every pixel has a number of short line segments (sticks) passing through it in all possible orientations. These sticks will be used as a template to find the most likely orientation along which a line segment exists. The sums of the pixel values along each line segment are evaluated and the pixel of interest takes the value of the maximum of the sums.

A compromise has to be made though, in terms of stick size selection, between speckle suppression and edge enhancement. Using a short stick means less reduction of speckle noise. On the other hand, a large stick increases the possibility of false edges being
detected. The authors found that sticks of length 15 yield a marked improvement in the
definition of the prostate boundaries.

To further suppress noise while preserving the edges the contrast enhanced image is
then smoothed by an algorithm called weak membrane fitting. The membrane is modeled
as a piecewise smooth function, which is associated with an energy. Minimising the energy
leads to a smoothed image with edge preserved.

Canny’s edge detector [223, 224] is then applied to the smoothed image to generate
an edge map. Next the user is required to specify the centre of the prostate. With the
designated centre as a reference point, the boundary points of the prostate are searched
in the radially outward direction. All the edge points corresponding to light-to-dark gray
scale transitions are identified as false edges and removed. The statistics of the remaining
edge points are collected in each quadrant with the centre of the prostate as a reference
point. In each quadrant the points which are far from the mean distance are removed.
The points that have survived will be presented as a visual guide for the user to manually
outline the prostate contour, which will be spline fitted and expanded by a constant factor
(20 pixels) to give a final smooth contour.

The authors claimed that this method is intended as an edge delineation method which
involves just-enough-interaction with the user. They suggested that having the flexibility
to accept or reject the edge segments detected is better than purely automatic approaches
in that patient-specific anatomical knowledge of the prostate can be incorporated through
the user interaction in the former. However, considerable work on manual edge linking
and editing is left to the user before a final meaningful contour can be formed.

8.3.4 Comments on the Edge Based Methods

Before any edge identification scheme is applied, the ultrasound images of the prostate
have to be filtered in the edge based strategies, due to low contrast and speckle noise.
In [265] a directional low-pass filtering method is used. Aarnink et al. [264] used a uniform
blurring filter because of its easy and fast application, whereas Pathak et al. [266, 267]
chose to use “sticks” to model edges as line segments. All these techniques have the same
consequence, that is, the edges are blurred and features smaller than the filter size are
displaced or merged with nearby features. Therefore the accuracy of the edges detected in
these smoothed image will be affected. Moreover, a compromise has to be made between
speckle reduction and edge localisation when it comes to the filter size, which is usually
selected on a trial and error basis. Once the filter size is decided, it remains unchanged
for all images of concern. There is no way to increase the accuracy of the edge detection
beyond the resolution set by the filtering process.

The edge detection procedures usually yield disrupted boundary segments rather than
a desired closed contour. Finding the correct edge segments that really represent part of
the prostate can be tricky and linking them and interpolating across the missing gaps to
form a closed contour is even more difficult. Automatic edge linking and interpolation can
be problematic [264] and this burden is completely transferred onto the user in [266, 267]. Using a centre based edge searching approach is a common characteristic among the edge based methods. The centre of the region that corresponds to the object of interest is indicated by the user. This centre is used as the starting point of a set of radial lines equally spaced in angle, along which candidate edge points will be searched for. This approach was adopted in other similar applications where the aim is to find the contour of an object whose shape resembles an oval or the like [277–281]. This idea has also made an impact on our contour extraction scheme based on the wavelet transform and snakes.

8.4 Model Based Strategies

The problems of interpolation for missing boundaries and edge linking with the edge based strategies can be totally avoided in the model based strategies [268–270], in which a priori information on the prostate shape is embedded in a template or model contour, and an effort is made to find the best match to the template in the image.

8.4.1 A Semi-Automatic Method Using Predefined Template Contours

Hendrikx and his co-workers [268] introduced a semi-automatic method for segmentation of ultrasound prostate images. The user has to identify a number of boundary points before the best fit of six predefined template contours is calculated to find out which contour these points are most likely to approximate.

The question raised here is whether the six template contours are representative enough for a wide range of prostate shapes in vivo. The answer is usually no and the best fit template contour is a rather crude representation of the real contour and large deviation from it can be expected. Another disadvantage of this method is the demanding human interaction to indicate the boundary points. The fitting process is also computationally expensive.

8.4.2 Ladak’s Model Based Initialisation

Ladak and his colleague [269, 270] developed a semi-automatic technique for prostate segmentation using a discrete dynamic contour (a numerical variant of the snake models) [282] and its model based initialisation. Four end points on the prostate boundary have to be indicated by the user beforehand: two in the antero-posterior direction and two in the latero-lateral direction. Hermitian interpolation of the four end points leads to an initial estimate of the prostate shape. This estimated contour is then used to initialise a discrete dynamic contour, which deforms automatically to fit the prostate contour.

The problem with this method is that the initialisation based on the four end points is not able to represent all possible variations of the prostate shape. The initial contour obtained after interpolation may lie far from the real prostate boundary, and/or going from bad to worse, close to extraneous edge points corresponding to surrounding tissues.
or other artifacts. For these cases, it will be trapped by local minima and cannot provide a satisfactory boundary delineation. Interactive guidance from the user is required to “push” the snake to better fit the prostate boundary.

8.4.3 Other Model Based Methods

There are other notable model based methods in the literature, including

- Knoll et al.’s shape restricted snakes [283,284] for segmentation of the prostate in CT and TRUS images. Inspired by the work of Chuang and Kuo [285,286], the authors proposed that the one-dimensional dyadic wavelet transform be used as a curve analysis tool, based on which the snake contour is parameterised and restricted. A multiresolution binary matched filter is applied to detect the object’s rough position and initialise the snake. Shape restriction is accomplished by replacing the wavelet coefficients of the contour (the so-called wavelet descriptors), which are significantly different from those of the most matched predefined models, with the corresponding model coefficients;

- Wu et al.’s method [287], which is based on feature (boundary) modelling and boundary searching with model constraint using an incremental genetic algorithm;

- Liu et al.'s method [288], which features preprocessing using an adaptive directional filter for noise removal and automatic attenuation compensation for contrast enhancement, a directional search strategy to locate key points on the prostate boundary, and interpolation from the key points under restriction of a morphological prostate boundary model; and

- recent work by Shen and colleagues [289], who use a Gabor filter bank within a statistical shape model of the prostate to segment the contour.

8.5 Final Remarks

In summary, the pixel or texture based methods are intrinsically deficient in differentiating prostate and non-prostate pixels. The model based strategies basically still rely, in one way or another, on the edge information derived from an edge detection algorithm or manual delineation. The model based strategies only fend off the difficulties of interpolation and edge linking which usually challenge and plague the pure edge based strategies.

The model based strategies rely, in one way or another, on a training phase, via which statistical models of the prostate boundary are established and derived from prior manual segmentation of a large number of images (which is a tedious and time-consuming task). During the automatic segmentation phase, optimisation in the parameter space of the statistical models is also computationally expensive. Usually the parameters of the statistical models require initialisation which can significantly affect performance of
these methods. Moreover, the large variations of the prostate shape \textit{in vivo} usually lands these strategies in trouble because the templates or predefined models are not necessarily representative.

Although the literature review in this chapter has been focused on a seemingly narrow range of papers, which specifically deal with the problem of prostate contour delineation, it is acknowledged that this problem shares many characteristics common to other medical imaging modalities and other tissue or organ structures, and thus is a general problem of segmentation of medical images \cite{290, 291}, MRI images \cite{292, 293} for instance. Obviously, there are no universal segmentation strategies that are applicable to medical images which will yield acceptable results in all cases due to peculiarity of various anatomical structures and imaging modalities. For example, there are no evident landmarks whatsoever in ultrasonic images of the prostate that can be taken advantage of for the purpose of segmentation, and ultrasonic images usually suffer from speckle noise and low-resolution. Moreover, our classification of the segmentation methods into the three classes may remain valid when examining other segmentation strategies of medical images of other anatomical structures or other modalities \cite{291}. Therefore, the “narrowness” of our literature review in this chapter is justifiable.

As a pilot study, the contour detection scheme based on the Laplacian of Gaussian (LoG) filter was introduced in Chapter 7. Basically it is an edge based approach, and inevitably has problems like those of similar approaches. One of the problems is that it is difficult to adaptively find a proper threshold. Too many or too few edges may be found depending on the threshold selected, which makes it difficult to track all the relevant edges that correspond to the prostate and form a closed contour from these boundary segments.

However, the experience of applying this scheme with LoG operators of different sizes suggested that it would be beneficial and viable to view the problem of segmenting the prostate contour from a multi-resolution perspective. The speckle noise reduction scheme in Chapter 6 also gives us such an insight. Our prostate contour extraction scheme is built exactly upon these perceptions, as described in Chapter 10.
Chapter 9

On Deformable Models

The last few decades have witnessed the rapid development of various noninvasive medical imaging technologies and modalities, such as computed tomography (CT), ultrasound, magnetic resonance imaging (MRI), positron emission tomography (PET), and emission computed tomography (ECT), just to name a few. They are becoming an increasingly important and sometime indispensable tool in routine clinical practice. With more and more medical images of different modalities flooding the usually overwhelmed radiologists’ offices every day, there is a growing need for automatic analysis and interpretation of the medical images by computers. Hence a wide spectrum of computer-assisted medical image analysis techniques are being investigated. Among them deformable models are one of the most promising and extensively researched areas, which range from the theoretically inclined algorithms to clinically motivated full-fledged systems.

In this chapter the theoretical background of deformable models will be introduced and their general applicability to medical image interpretation will be outlined to lay the foundations for the contour extraction scheme to be described in Chapter 10. This chapter is organised as follows. First the physically motivated energy minimisation formulation of traditional snake models will be introduced. Minimising the snake energy functional is a variational problem and Euler’s equation for snakes is presented in 9.3 and numerical simulation is described in 9.4. More generally, the deformable models can be formulated on the basis of the principles of Lagrangian mechanics, which leads to dynamic deformable models. These models make a valuable tool in medical image sequence analysis such as motion estimation and tracking and will be introduced in 9.5. The traditional snake models have their own pitfalls, such as sensitivity to initial conditions, limited attraction range, and poor convergence to concavities. More external forces and several variations and extensions to the traditional parametric deformable models will be introduced in 9.6 and 9.7 respectively. Finally in the last section, the deformable models and their various issues will be summarised and discussed. Their implications for the medical image analysis in general and the contour extraction scheme in Chapter 10 in particular will be described.
9.1 Introduction

“Deformable models are a powerful, physics-based technique for representing, reconstructing, recognizing, and manipulating nonrigid curves, surfaces, and solids from their images and image sequences” [294]. A snake\(^1\) is an energy-minimizing curve within an image domain. It is constrained by its inherent internal forces and guided by external image forces that pull it towards prominent image features such as lines and edges.

Snakes or active contour models were first introduced by Kass, Witkin and Terzopoulos in their seminal paper “Snakes: active contour models” [174,295]. However, the idea of using physics-motivated models of non-rigid objects to extract image features can be traced back to Fischler and Elshlager’s spring-loaded templates [296], whose components are held together by spring forces, and Widrow’s rubber mask technique [297], in which parameterised templates called rubber masks are used to model object of interest. Similar ideas of detecting edges and curves by optimising models of curve contrast and smoothness can also be found in the work of Martelli [298,299] and Montanari [300]. Deformable models generalise these ideas through the use of elastodynamic models and by incorporating internal, external and interactive guiding forces in an energy minimizing framework.

Since the publication of the paper by Kass et al., deformable models have enjoyed fast-growing popularity within the computer vision community and especially in the area of medical image interpretation [301–309]. This popularity is largely due to their unique features.

1. Deformable models simulate elastic material which can dynamically conform to objects in response to internal forces, external image forces, and user specified constraints. The most prominent feature of deformable models, which differentiates them from previous, purely geometric models, is the introduction of an energy function by analogy with that of physical systems. By manipulating the energy function, a wide range of deformable models can be specified to suit different image processing scenarios.

2. Through the energy function, shape constraints such as continuity, smoothness, symmetry and curvature are transformed into intrinsic internal forces. Likewise, external forces may be derived from image data or interactive user manipulation. Deformable models respond to these internal and external forces just as objects respond to applied forces in the physical world. Under the influence of these simulated forces, they evolve to conform to image structures of interest, such as lines, edges, and subjective contours. The interactive or automatic adjustment of deformable models with these forces involved is therefore very natural and intuitive.

\(^1\)Deformable models have also been referred to as snakes or active contour models in the literature. These terms are interchangeable in the following text, although rigorously speaking snakes are only a two-dimensional special case of deformable models, which is a general formalism. Since we will only focus on (two-dimensional) images in the following text, this will not cause any confusion.
3. Deformable models are fundamentally dynamic, unifying the analysis of shape and motion by combining geometry and physics. This makes them especially amenable to estimating motion to track moving natural objects in image sequences, such as beating hearts, moving lips, articulating limbs, and so on.

The deformable models provide a unified treatment of a wide range of image interpretation problems that have been addressed with different approaches before. Important image features such as edges, lines and subjective contours can all be easily handled in the same framework. Tracking these features through motion and matching them in stereo can also be treated by essentially the same mechanisms. Because of these virtues, deformable models have found wide application in image segmentation, motion tracking and analysis, image matching and 3D reconstruction.

9.2 Energy Minimising Formulation

A traditional snake [174, 295] is a planar parametric curve \( \mathbf{v}(s) = (x(s), y(s)) \) which moves within the image domain \( (x, y) \in \mathbb{R}^2 \) minimising its energy, which is defined as the sum of internal energy and external potential energy,

\[
E_{\text{snake}}(\mathbf{v}) = E_{\text{int}}(\mathbf{v}) + E_{\text{ext}}(\mathbf{v}) \tag{9.1}
\]

where \( \mathbf{v} \) represents the curve as a mapping from the unit parametric domain \( s \in [0, 1] \) to the image plane \( \mathbb{R}^2 \). The components of \( \mathbf{v}(s), (x(s), y(s)) \), are the curve’s coordinate functions. The internal energy is designed to control the smoothness and the tension of the curve, and is transformed into internal forces, which smooth the snake or make it shrink during its deformation. The external energy is expected to fit the snake contour to intensity extrema, boundaries and other features of interest within an image. The external energy is defined on the basis of a scalar potential function, which is derived from the image so that its local minima coincide with the image features of interest. By specifying an appropriate energy functional, a snake can be “customised” to possess particular properties to suit different image processing needs.

9.2.1 The Internal Energy

The internal energy of a snake, represented by the first term on the right-hand side of (9.1), depends on intrinsic properties of the snake, such as its continuity and curvature. It is defined as the sum of the of elastic energy (of relevance to its continuity) and bending energy (affecting its curvature),

\[
E_{\text{int}} = \int_0^1 \left( \frac{1}{2}(\alpha(s)|\mathbf{v}'(s)|^2 + \beta(s)|\mathbf{v}''(s)|^2 \right) ds \tag{9.2}
\]
where $v'$ and $v''$ respectively denote the first and second derivatives of $v$ with respect to $s$. The first term in the integral in (9.2) is the elastic energy (or energy of continuity) and the second term is the bending energy (or energy of curvature).

The physical behaviour of the snake, such as its "tension" and "rigidity", and properties like its local continuity and curvature, are controlled by its two non-negative regularisation parameter functions $\alpha(s)$ and $\beta(s)$:

- $\alpha(s)$ controls the "tension" of the snake and makes it behave more like an elastic string obeying Hooke’s law. A large $\alpha$ increases the "tension", discourages stretching and gives a dominant effect of shrinking or contracting. So the snake tends to reduce its length resulting in extraneous loops and ripples being eliminated. Unlike most real world springs, which have a finite resting length, a snake would eventually evolve to shrink to zero length if it was under sole influence of the elastic energy;

- $\beta(s)$ controls the "rigidity" of the curve and makes it behaves more like a thin rigid (but still flexible) metal strip. A large $\beta$ increases the “rigidity” of the snake and discourages bending, making the snake less flexible and giving a dominant effect of smoothing. If $\beta$ is set to zero at any point $v_i$ on the snake, that is, the influence of the bending energy is switched off, a sharp corner is allowed to form at $v_i$. In particular, if $\alpha = 0$ and $\beta = 0$ at $v_i$, then a position discontinuity is permitted there.

The effect of the snake regularisation parameters is demonstrated in Figure 9.1. Figure 9.1(a) shows a snake with 28 control points in its initial state. The snake is made to evolve for 100 iterations with different regularisation parameters. The deformation of the snake is displayed every 5 iterations and the results are shown in Figure 9.1(b), (c) and (d). As we can see in Figure 9.1(b), the snake initialised with regularisation parameters $\alpha = 0.1$, $\beta = 0$ keeps contracting into a smooth circle. The control points are quickly pulled in-line with their neighbours by the elastic forces. The dominant effect in this case is shrinking, although this is coupled with the effect of smoothing. In Figure 9.1(c) the regularization parameters are set to $\alpha = 0$, $\beta = 0.1$. The evolving snake trajectories show that the snake is straightened out by the bending forces and the dominant effect in this case is smoothing rather than shrinkage.

The effect of changing the parameters $\alpha, \beta$ is more evident when Figure 9.1(b) is compared with Figure 9.1(d), where $\alpha = 0.1$, $\beta = 0.1$. As we can see, $\beta$ plays a more important role than that of $\alpha$ at the early stage of the snake deformation to straighten out its ragged contour. Thanks to the bending force the snake is smoothed more quickly than it is in Figure 9.1(b), where the bending force is lacking. After the snake is smooth enough, its deformation will be dominated by $\alpha$, which makes it keep contracting. This explains why the later stages of the snake deformation in Figure 9.1(b) and Figure 9.1(d) are quite comparable, although the snake in Figure 9.1(d) is collapsing a little bit faster than that in Figure 9.1(b) due to presence of the bending force as well as the elastic force.
ON DEFORMABLE MODELS

9.2.2 The External Energy

The second term on the right-hand side of (9.1) is an external energy term which couples the snake to image features via a scalar potential function $P(x, y)$ typically derived from raw image data $I(x, y)$

$$E_{\text{ext}} = \int_0^1 P(v(s))ds$$  \hspace{1cm} (9.3)

The potential function $P(x, y)$ generally is designed to take smaller values at object boundaries or other image features of interest. A snake’s behaviour can then be further manipu-
lated by designating different external energy functionals which attract a snake to salient features in the image, such as lines and edges.

The simplest form of the potential energy function is the image intensity itself. If the image features bright lines on a dark background or dark lines on a light background, the external energy function can be set as

\[ P(x, y) = -\gamma I(x, y) \]  

(9.4)

Depending on the sign of the weighting parameter \( \gamma \), the snake will latch on to the bright lines (\( \gamma > 0 \)) or dark lines (\( \gamma < 0 \)) nearby.

If the image contains bright regions (rather than lines) on a dark background (or vice versa), the boundaries can be found by a snake with the external energy function

\[ P(x, y) = -\gamma |\nabla I(x, y)| \]  

(9.5)

where \( \nabla \) is the gradient operator. The snake will try to align itself with the contours with large amplitude of gradients, which correspond to the edges of the bright regions of interest. Apparently \( \gamma \) takes positive values here.

The influence of the external energy is illustrated in Figure 9.2 where a snake is placed within an image which has gradient pointing towards its centre, as shown in Figure 9.2(b). Here the external potential function is set as in equation 9.4. The snake deforms for 100 iterations with different regularisation parameters in each case. In Figure 9.2(c) the influence of the internal energy of the snake is totally switched off by setting internal energy parameters \( \alpha = 0.0, \beta = 0.0 \). The external energy weighting parameter is set to \( \gamma = 0.5 \). As we can see in Figure 9.2(c) the snake is “squeezed” towards the centre portion of the image where it is brighter, while relatively maintaining its original shape. Should \( \gamma \) be negative, the snake would swell into the dark portions of the image.

Both the internal energy and the external potential energy are allowed to come into play in Figure 9.2(d), where \( \alpha = 0.1, \beta = 0.1, \gamma = 0.5 \). Due to the external forces, which are pointing inwards, the snake is pushed into the centre of the image faster compared to the snakes in Figure 9.1(b) and Figure 9.1(d). Should \( \gamma \) be negative, the collapsing of the snake would either be slowed down or turned into expansion, depending on the magnitudes of \( \alpha, \beta \) and \( \gamma \) and their mixed effect.

It is not uncommon for many real world images to contain lots of noise or small-scale spurious local structures. In these cases the evolution of the snake will be misled or disrupted by the random forces generated by these disturbances. To circumvent these disturbances the image can be smoothed to filter out the noise and small features to make the snake “see” more global structures. A Gaussian function, which is optimally localised in both space and frequency domain, is usually used as the smoothing filter.
Figure 9.2: The influence of the external energy. (a) A snake with 28 control points put in an image with gradient pointing inwards. (b) The gradient of the image. (c) After 100 iterations, $\alpha = 0$, $\beta = 0$, $\gamma = 0.5$. The snake is “squeezed” towards the centre of the image while maintaining its shape. (d) After 100 iterations, $\alpha = 0.1$, $\beta = 0.1$, $\gamma = 0.5$. The snake is contracting faster due to the external force.

For the above two cases, the external energy function can also be set to

$$P(x, y) = -\gamma G_{\sigma}(x, y) * I(x, y)$$  \hspace{1cm} (9.6)

or

$$P(x, y) = -\gamma |\nabla [G_{\sigma}(x, y) * I(x, y)]|$$  \hspace{1cm} (9.7)

where $G_{\sigma}(x, y)$ is a two-dimensional Gaussian function with standard deviation $\sigma$ and $*$ denotes the convolution operation. A large $\sigma$ is often necessary for a snake to increase its
9.3 Euler’s Equation for Snakes

Substitution of (9.2) and (9.3) into (9.1) gives the snake energy functional

$$E_{\text{snake}} = \int_{0}^{1} \left( \frac{1}{2} (\alpha(s)|v'(s)|^2 + \beta(s)|v''(s)|^2) + P(v(s)) \right) ds.$$  (9.8)

Determining the minimum of the energy functional of snakes falls into a wide class of problems that is the subject of the Calculus of Variations [310, 311], where the objective is to determine extrema or stationary values for functionals. It is known that the extrema or stationary values for a functional must satisfy the corresponding Euler’s equation as a necessary condition (see Appendix A). For the case of the snakes problem, Euler’s equation is

$$\frac{d}{ds} (\alpha v') - \frac{d^2}{ds^2} (\beta v'') - \nabla P(v) = 0$$  (9.9)

In practice both $\alpha$ and $\beta$ are constants, and in this case the above reduces to

$$\alpha v'' - \beta v'''' - \nabla P(v) = 0$$  (9.10)

which cannot be explicitly integrated and thus a direct solution cannot be obtained. Numerical methods are usually employed to find a solution to (9.9) (see next section).

Equation (9.9) can also be interpreted from a different point of view. Here the physical analogy can be applied again. Just as the forces on physical systems always make them move so as to reduce (or increase) their energy, a snake reducing its total energy can be described by some simulated forces applied to the snake. Now the motion of the snake becomes physically intuitive and “legitimate” due to these simulated forces acting on it. When the snake reaches its lowest energy state and comes to rest, (9.9) expresses this equilibrium state as the balance of these forces.

$$\mathbf{F}_{\text{int}} + \mathbf{F}_{\text{ext}} = 0$$  (9.11)

where $\mathbf{F}_{\text{int}}$ denotes the internal force stemming from the internal energy of the snake

$$\mathbf{F}_{\text{int}} = \frac{d}{ds} (\alpha v') - \frac{d^2}{ds^2} (\beta v'')$$  (9.12)

and the negative gradient of the image potential is interpreted physically as the external
potential force corresponding to the external potential energy of the snake

\[ F_{\text{ext}} = -\nabla P(v) \]  

(9.13)

The first term in (9.12) is the elastic force which encourages the snake’s tendency to contract or shrink. The second term is the bending force which discourages sharp corners in the curve. The external force drives the curve toward the salient features within the image, which usually possess a large amplitude of gradient. The external forces are usually referred to as Gaussian potential forces if the corresponding external energy functions are defined by equations (9.6) and (9.7).

In addition to image structures the external energy can also depend on factors such as particular constraints the user has imposed. More external forces will be introduced in section 9.6.

9.4 Numerical Simulation

Since the snake model was proposed, various numerical simulation approaches have been investigated, including the finite difference method [174], the finite element method [303, 304, 312–314], dynamic programming [315], a discrete dynamic contour model [282], and a greedy algorithm [316], just to name a few. The finite difference method is efficient to compute. The finite element method is more expensive computationally but on the other hand has the advantage of being easily adapted to the irregular mesh representation of deformable surfaces. In this section we focus on the finite difference method [174].

In order to numerically compute a solution to (9.8), the snake is made dynamic by treating \( v(s) \) as a function of time \( t \) as well as \( s \), that is, \( v(t, s) \). The product of a step size \( \eta \) and the partial derivative of \( v \) with respect to \( t \) is then set equal to the left hand side of (9.9) as follows

\[ \eta \frac{\partial v}{\partial t} = \frac{d}{ds} (\alpha v') - \frac{d^2}{ds^2} (\beta v'') - \nabla P(v) \]  

(9.14)

When the time derivative (left-hand side of (9.14)) vanishes, \( v(t, s) \) stabilises and we obtain a solution to (9.8). This approach of making the time derivative term vanish is equivalent to applying a gradient descent algorithm to locate the local minimum of (9.8) [313,317].

Next it is necessary to discretise the energy functional \( E_{\text{snake}} \). A general approach is to represent the argument function \( v \) on a set of \( N \) discrete nodes \( v_i = v(ih) \) for \( i = 0, \ldots, N - 1 \) where \( N \) is the total number of nodes and \( h = 1/(N - 1) \) is the step size in space. The derivatives are then approximated by finite differences as

\[ v'_i \approx (v_i - v_{i-1})/h \]
\[ v''_i \approx (v_{i+1} - 2v_i + v_{i-1})/h^2 \]  

(9.15)
Now (9.14) can be rewritten as

\[
\eta_i \frac{v_i^n - v_i^{n-1}}{\Delta t} = \frac{1}{h^2} \left[ \alpha_{i+1}(v_{i+1}^n - v_i^n) - \alpha_i(v_i^n - v_{i-1}^n) \right] \\
- \frac{1}{h^4} \left[ \beta_{i+1}(v_{i+2}^n - 2v_{i+1}^n + v_i^n) \\
- 2\beta_{i}(v_{i+1}^n - 2v_i^n + v_{i-1}^n) \\
+ \beta_{i-1}(v_i^n - 2v_{i-1}^n + v_{i-2}^n) \right] \\
+ F_{ext}(v_i^{n-1})
\]  

(9.16)

where \( v_i^n = (x(\alpha_i, n\Delta t), y(\beta_i, n\Delta t)) \). \( \eta \) is also referred to as damping coefficient or damping density. \( \alpha_i = \alpha(\alpha_i) \) and \( \beta_i = \beta(\beta_i) \) are sampled at the same nodes as \( v_i \). \( F_{ext}(v_i) \) is a vector of the discrete version of the external force.

Equation (9.16) defines \( N \) simultaneous equations and they can be assembled into a compact matrix form as

\[
\frac{1}{\kappa}(v^n - v^{n-1}) = K v^n + F_{ext}(v^{n-1})
\]

(9.17)

where \( \kappa = \Delta t/\eta \) is the iteration step size and \( K \) is a \( N \times N \) symmetric pentadiagonal banded matrix

\[
K = \begin{bmatrix}
& & & & & c_{N-2} & b_{N-1} \\
& & & & & a_0 & b_0 & c_0 \\
& & & & b_0 & a_1 & b_1 & c_1 \\
& & & c_0 & b_1 & a_2 & b_2 & c_2 \\
& & c_1 & b_2 & a_3 & b_3 & c_3 & \cdots \\
& c_N & c_{N-1} & c_{N-2} & a_{N-3} & b_{N-3} & c_{N-3} & \cdots \\
& c_{N-4} & c_{N-3} & c_{N-2} & a_{N-2} & b_{N-2} & \cdots \\
& c_{N-4} & c_{N-3} & c_{N-2} & a_{N-1} & b_{N-1} & \cdots
\end{bmatrix}
\]

(9.18)

where the unspecified entries are all 0’s and

\[
a_i = (\alpha_i + \alpha_{i+1})/h^2 + (\beta_{i-1} + 4\beta_i + \beta_{i+1})/h^4 \\
b_i = -\alpha_{i+1}/h^2 - 2(\beta_i + \beta_{i+1})/h^4 \\
c_i = \beta_{i+1}/h^4
\]

(9.19)

Cyclic boundary conditions, that is, a closed contour, are assumed here.

Equation (9.17) can then be solved iteratively by the matrix operation

\[
v^n = (I - \kappa K)^{-1} \left[ v^{n-1} + \kappa F_{ext}(v^{n-1}) \right]
\]

(9.20)

\((I - \kappa K)\) is also a symmetric pentadiagonal banded matrix. Its inverse can be obtained efficiently by \( LU \) decomposition in \( O(N) \) complexity [174] and only needs to be done
ON DEFORMABLE MODELS

once because the regularisation parameters for each node remain unchanged during the iterative processes. Hence (9.20) provides a fast numerical solution to (9.9) through matrix manipulation.

9.5 Dynamic Deformable Models

In section 9.3 the energy minimisation of deformable models was considered naturally as a static problem. In order to find a numerical solution to the energy functional (9.8), a snake was artificially made dynamic in section 9.4, by virtue of a new variable of time \( t \) being introduced for \( \mathbf{v}(s) \).

More generally, however, it is sometimes more convenient or necessary to formulate the deformable models on the basis of the principles of Lagrangian mechanics by constructing a dynamical system that is governed by the energy functional and allowing the system to evolve to a minimal energy state as it achieves equilibrium. Such a formulation leads to dynamic deformable models that unify the description of both shape and motion, making them sensible of the analysis of not only static shapes but also shapes that evolve through time. This makes them especially valuable for medical image sequence analysis, since most natural objects, such as beating hearts, articulating limbs and so on, are non-rigid and continually undergo motion \textit{in vivo}. Thus dynamic deformable models make a valuable tool in motion estimation and tracking.

Moreover, dynamic deformable models offer a variety of intuitively meaningful physical behaviours that are not necessarily evident from the static energy minimisation point of view. For example, a dynamic snake may be guided interactively by the user as it minimises its energy, making its evolution amenable to the user’s intervention.

9.5.1 The Lagrangian Formulation of Dynamic Deformable Models

The Lagrangian formulation of dynamic snakes was first proposed by Terzopoulos [318, 319]. A dynamic snake is represented by introducing a time-varying mapping \( \mathbf{v}(s,t) \) and a kinetic energy

\[
E_k = \int_0^1 \frac{1}{2} \mu(s) |\dot{\mathbf{v}}(s,t)|^2 ds
\]

(9.21)

where \( \mu(s) \) is the mass density, and \( \dot{\mathbf{v}} \) denotes the first derivative of \( \mathbf{v} \) with respect to time \( t \). The kinetic energy and the deformation potential energy in (9.8) are combined to give the energy difference, which is generally called the \textit{kinetic potential} or \textit{Lagrangian function}

\[
\mathcal{L}(\mathbf{v}) = E_k - E_{\text{snake}}
= \int_0^1 \frac{1}{2} \mu(s) |\dot{\mathbf{v}}|^2 - \frac{1}{2} \left( \alpha |\mathbf{v}'|^2 + \beta |\mathbf{v}''|^2 \right) - P(\mathbf{v}) ds
\]

(9.22)
9.5. DYNAMIC DEFORMABLE MODELS

According to Hamilton’s variational principle (Appendix B), between two instants of time \( t_0 \) and \( t_1 \), the dynamic deformable model’s motion from its initial state \( \mathbf{v}(s, t_0) \) to its final equilibrium state \( \mathbf{v}(s, t_1) \) proceeds in such a way that the time integral of the difference between the kinetic and potential energies

\[
J = \int_{t_0}^{t_1} \mathcal{L} (\mathbf{v}) dt
= \int_{t_0}^{t_1} \int_0^1 \frac{1}{2} \mu |\dot{\mathbf{v}}|^2 - \frac{1}{2} \left( \alpha |\mathbf{v}'|^2 + \beta |\mathbf{v}''|^2 \right) - P(\mathbf{v}) ds dt
\] (9.23)

will be stationary, that is, its variation with respect to \( \mathbf{v} \) vanishes

\[
\delta J = \delta \int_{t_0}^{t_1} \mathcal{L} (\mathbf{v}) dt = 0 \] (9.24)

This condition leads to Lagrange’s equations of motion for the model

\[
\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{v}}} \right) - \frac{\partial \mathcal{L}}{\partial \mathbf{v}} + \frac{\partial}{\partial s} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{v}'} \right) - \frac{\partial}{\partial s^2} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{v}''} \right) = 0 \] (9.25)

with appropriate initial and boundary conditions, where \( \mathcal{L} \) represents the integrand in (9.23).

This gives rise to another problem. Once in motion, unless its kinetic energy is somehow dissipated, such a dynamic snake with a mass distribution will move perpetually, making it impossible to achieve static equilibrium. To address this problem a Rayleigh dissipation functional

\[
\mathcal{D}(\dot{\mathbf{v}}) = \int_0^1 \xi(s) |\dot{\mathbf{v}}|^2 ds \] (9.26)

is introduced and incorporated into the model to dampen the snake, where \( \xi(s) \) is the damping density. After the variational derivative of the integrand \( \mathcal{D} \) in the dissipation functional (9.26) is included, Lagrange’s equation of motion becomes

\[
\frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{v}}} \right) + \frac{\partial \mathcal{D}}{\partial \dot{\mathbf{v}}} - \frac{\partial \mathcal{L}}{\partial \mathbf{v}} + \frac{\partial}{\partial s} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{v}'} \right) - \frac{\partial}{\partial s^2} \left( \frac{\partial \mathcal{L}}{\partial \mathbf{v}''} \right) = 0 \] (9.27)

In practice mass density \( \mu(s) \) and damping density \( \xi(s) \) are set as constants and the equation above reduces to

\[
\mu \ddot{\mathbf{v}} + \xi \dot{\mathbf{v}} - \frac{\partial}{\partial s} (\alpha \mathbf{v}') + \frac{\partial^2}{\partial s^2} (\beta \mathbf{v}'') = -\nabla P(\mathbf{v}) \] (9.28)

This partial differential equation suggests a straightforward force balance mechanism. The first two terms on the left-hand side of (9.28) denote inertial and damping forces respectively. From (9.12), apparently the last two terms on the left-hand side represent the internal forces which consist of the elastic force and bending force. These forces balance the generalised external forces on the right-hand side. When an equilibrium is achieved
and the snake comes to rest, the derivatives with respect to time vanish, $\dot{v} = \ddot{v} = 0$, leading to the static solution (9.9) for the energy minimising formulation of deformable models which indicates that it is a special case of the dynamic framework.

Equation (9.28) can also be interpreted from a different point of view, that is, Newton’s second law of motion. The fact that Lagrange’s equations of motion imply Newton’s second law makes this interpretation legitimate. According to Newton’s second law, the time rate of change of the momentum of a body equals the net forces exerted on it, and equation (9.28) can be written as

$$\mu \ddot{v} = F_d + F_{int} + F_{ext}$$

(9.29)

where

$$F_d = -\xi \dot{v}$$

(9.30)

denotes the damping force and

$$F_{int} = \frac{\partial}{\partial s} (\alpha v') - \frac{\partial^2}{\partial s^2} (\beta v'')$$

(9.31)

is the internal elastic and bending forces like that defined in equation (9.12). $F_{ext}$ is the external potential force given in equation (9.13).

One way to visualise the dynamic deformable model is to think of the image potential $P(x, y)$ as an imaginary terrain with mountain peaks, plateaus (corresponding to high magnitude of the image potential), valleys and ravines (low magnitude of the image potential). A snake in the terrain tends to slide downhill from shoulders and ridges to the ravines and valleys, conforming to the contour lines while on its way.

Dynamic deformable models can also be applied to sequences of images. For an image sequence, the potential $P(x, y, t)$ will be time-varying. Suppose a snake has achieved equilibrium for frame $k$ where image potential is $P(x, y, t_k)$. The potential $P(x, y, t_{k+1})$ of the next frame $k + 1$ is a perturbation of $P(x, y, t_k)$ and the valleys or ravines where the snake rests will have shifted. The potential energy is converted to kinetic energy and the snake will evolve non-rigidly towards and finally conform to the shifted ravines to reach a new equilibrium.

### 9.5.2 Numerical Simulation

To simulate the dynamic snakes and find a numerical solution to (9.28), we need to deal with the discretised version of (9.28), which is a series of second order ordinary differential equations for the discretised vector $v(t)$

$$M \ddot{v} + C \dot{v} + K v = F$$

(9.32)

where $K$ is called the stiffness matrix as defined in (9.18). $M$ is a mass matrix and $C$ is a damping matrix. Both of them are sparse banded matrices like $K$. $F$ is a vector of the discrete version of the generalised external force.
9.6 External Forces

This system of ordinary differential equations must be integrated forward through time and several explicit and implicit time-integration methods can be applied. A semi-implicit method is demonstrated here, for which the time step is supposed to be $\Delta t$. Thus the derivatives of $v$ with respect to time $t$ are replaced with backward finite difference

\[ \ddot{v} \approx \frac{(v[t+\Delta t] - 2v[t] + v[t-\Delta t])}{(\Delta t)^2} \]  

and

\[ \dot{v} \approx \frac{(v[t+\Delta t] - v[t-\Delta t])}{2\Delta t} \]  

where the superscripts denote that the variables concerned are evaluated at the time given in brackets. The update formula is given by

\[ Av[t+\Delta t] = b[t] \]  

where

\[ A = M/(\Delta t)^2 + C/2\Delta t + K \]  

is symmetric and banded and

\[ b[t] = (2M/(\Delta t)^2)v[t] + (C/2\Delta t - M/(\Delta t)^2)v[t-\Delta t] + F[t]. \]  

The pentadiagonal matrix $A$ can be efficiently factorised into into lower and upper triangular matrices as

\[ A = LDL^T \]  

where $L$ is a lower triangular matrix and $D$ is a diagonal matrix. Since $A$ is constant in practice, the LU factorisation is unique and only needs to be done once. The solution $v[t+\Delta t]$ is then obtained by solving $LS = b[t]$ with forward substitution and then solving $L^Tv = D^{-1}S$ with backward substitution.

9.6 External Forces

The external forces derived from the external potential energy function $P(x, y)$ defined in (9.4) and (9.5) are potential forces. The external forces corresponding to the external potential energy function $P(x, y)$ defined in (9.6) and (9.7) are called Gaussian potential forces. All these forces have a common characteristic—they are defined as the negative gradient of a potential energy function. It is not physically intuitive to incorporate non-potential forces into the traditional snake model (section 9.2) since these forces cannot be derived from the energy minimisation formulation. The dynamic formulation discussed in the previous section permits the incorporation of generalised external forces, both potential and non-potential external forces. More external forces will be introduced in this section.
Interactive Constraint Forces

Although a snake has typically been regarded as “autonomous” after initialisation, it is desirable to interact with it during its deformation to guide it towards an appropriate local minimum. After all, deformable models have been designed as an interactive tool for semi-automatic image interpretation from the very beginning. This kind of interaction becomes more important in some situations when the inherent forces on the snake fail to make it land onto the desired features in certain regions where the interference from spurious local structures is dominant, or when the user would like to intervene in the deformation of the snake so that it must pass through some landmark points specified by the user.

Two different interactive constraint forces have been reported in the literature, spring forces and repulsion forces [174, 295, 312]:

**Spring forces** are designed to pull a snake towards desired image features or a “pin” point specified by a user. They are implemented to be proportional to the separation between a pin point \( p \) and the point of influence \( v_i \) on the model’s contour

\[
F_s = \kappa_s (p - v_i) \tag{9.39}
\]

The further away the model is from the pin point \( p \), the stronger the spring force, whose strength can be further fine-tuned by the coefficient \( \kappa_s \). The point of influence \( v_i \) on the model’s contour can be identified as the nearest node to the pin point \( p \) using a heuristic local neighbourhood search.

**Repulsion forces** are, on the contrary, used to push a snake out of one local minimum and into another. Repulsion forces can be defined as

\[
F_r = \kappa_r \frac{r}{|r|^2} \tag{9.40}
\]

where \( r = p - v_i \). Note that the repulsion forces must be “clipped” (set to zero) near \( r = 0 \) to prevent numerical instability, resulting in that the force “landscape” looks like a volcano. Hence these forces are also called volcano forces.

These interactive forces allow a user to manipulate the deformation of the snakes and to make them have few local minima and little dependence on starting points.

Inflation Force

The traditional deformable models have their own limitations. One of them is that snakes are sensitive to their initial conditions. In other words, their “capture” range is quite limited and the models must be initialised close to the desired object boundaries. Moreover, due to noise, some isolated spurious edge points may be gradient maxima and can stop the
9.6. EXTERNAL FORCES

model when it approaches. To address this problem an outward inflation force was introduced by Cohen [304], which is expected to push a snake outward as if air were introduced inside the model, increasing the attraction range of the snake and making the model pass by isolated spurious edge points and less sensitive to initial conditions. Active contour models that utilise inflation force are therefore also referred to as balloons [303, 304, 313, 320].

The inflation force of a balloon is defined as

\[ F_b = \kappa_b \vec{n}(v) \]  \hspace{1cm} (9.41)

where \( \vec{n}(v) \) is the unit vector normal to the model contour at point \( v \) and the coefficient \( \kappa_b \) can be regarded as the amplitude of the inflation force. \( \kappa_b \) is typically chosen by the user such that it is of the same order as of the potential force weighting parameter \( \gamma \) in equations (9.4) to (9.7) and slightly smaller than \( \gamma \) so that an edge point of large amplitude can balance the inflation force and prevent further expanding of the model.

A balloon is usually placed in the interior of a desired object contour of interest. It keeps expanding and is attracted and temporarily stopped by edges nearby like a traditional snake model. Because of the presence of the inflation force, however, if this edge is too weak to make the model stick to it, the model will pass through this edge and continue expanding until it encounters a strong edge segment which is strong enough to balance the inflation force to make it reach a new equilibrium.

The sign of \( \kappa_b \) can be flipped to introduce a deflation force rather than an inflation force. A deformable model like this is usually initialised outside the object of interest and shrinks and finally locks on to the object.

The inflation or deflation forces are powerful “assistants” to make the deformable models converge to the desired solution being less demanding of the initialisation of the model. A disadvantage of introducing the inflation force or deflation force in the deformable models is, however, they may make the contour cross itself and form loops, which is not desirable [317, 321].

9.6.3 Distant Attraction Potential Force

In an attempt to combine the advantages of a good local edge detector with a global active contour model, the distant attraction force was proposed by Cohen and Cohen [320], using a distance map to extend the capture range of the snakes.

The distant attraction force must be defined through the use of a potential function, which may be defined as the convolution of a binary edge map with a Gaussian function. The potential function can also be defined using a distance map, which is obtained as follows. First a binary edge map of the original image is obtained using a local edge detector, such as Canny’s edge detector [223, 224]. Then the distance map is calculated by finding the distance between each pixel and the nearest edge point, either using Euclidean distance [322] or using Chamfer distance [303, 323] to approximate the Euclidean distance.
Let the distance map denoted by \( d(x,y) \). A variety of potentials can be defined as a function of the distance map \( d(x,y) \), formulated as \( P(x,y) = G(d(x,y)) \). For example,

\[
P(x,y) = -e^{-d(x,y)^2}
\]

(9.42)
gives a potential which is similar to the Gaussian convolution mentioned above. It can also be defined as

\[
P(x,y) = \begin{cases} 
-1 & d(x,y) \leq 1 \\
\frac{1}{d(x,y)} & d(x,y) > 1 
\end{cases}
\]

(9.43)

Thus the corresponding potential force for these cases becomes

\[
F_p = -\nabla P = -G'(d)\nabla d
\]

(9.44)

To make this force independent of function \( G \) it was suggested to be normalised as

\[
F_p = -\kappa_p \frac{\nabla d}{||\nabla d||}
\]

(9.45)

The distant attraction force derived from these modified potentials gives a faster convergence to the final solution since these potentials decay slowly, producing larger forces at points far from the edge points and increasing the capture range of the snakes. The distant attraction force may be used as a replacement of the traditional potential force corresponding to equations (9.4) to (9.7) or combined with them to enhance edge detection. The latter is especially useful when dealing with disconnected edge segments [320].

9.7 Variation and Extension of the Original Snake Model

The traditional parametric deformable model has spawned many notable variations and extensions including probabilistic deformable models [324–326], Fourier deformable models [327] and gradient vector flow [175,176,317,328]. These variations and extensions were intended to improve the traditional parametric deformable models in the following aspects:

- **Incorporation of prior information.** Although the traditional parametric snake model was designed as a “global” boundary finding method, there is no global shape constraint or other *a priori* information whatsoever built into the model. The advantage of incorporating *a priori* information into the deformable models is that it makes them immune to the interference from the spurious local image structures and more robust results can be expected. This can be accomplished either by casting the problem of energy minimisation in a probabilistic framework, as in probabilistic deformable models, or through the use of global shape constraint, as in Fourier deformable models. All of them involve a training process from which prior information on the variability of the object of interest are collected from some sample images.
9.7. VARIATION AND EXTENSION OF THE ORIGINAL SNAKE MODEL

- **Sensitivity to initialisation.** The capture range of a snake is limited. A snake must be initialised around the immediate vicinity of the desired object. Otherwise, it will not be attracted by the object and might be misled by the nearby spurious features. Although the balloon model and distant attraction potential force have been proposed to address this problem, as described in previous sections, they have their own limitations.

- **Poor convergence to boundary concavities.** A traditional snake usually experiences difficulty in deforming towards regions with concave boundaries. No satisfactory solution to this problem had been found until the proposition of gradient vector flow snake, which is expected to address the initialisation problem of the snake model as well.

### 9.7.1 Probabilistic Deformable Models

The energy minimisation form of the traditional parametric deformable models stems from physical analogy, which is natural and intuitive, as described in section 9.2. It is also possible, however, to establish an alternative to this physical point of view by putting the problem of energy minimisation in a probabilistic framework and viewing it as an estimation problem [324–326]. The probabilistic framework makes it feasible to incorporate into the deformable models a priori information and sensor model characteristics in terms of probability distributions. Deformable models may then be interpreted by virtue of Bayesian estimation.

Suppose \( v \) denotes the unknown deformable model shape parameters. According to Bayes’ Theorem [329], the posterior distribution of \( v \) given the image data \( I \) is

\[
p(v|I) = \frac{p(I|v)p(v)}{p(I)} \tag{9.46}
\]

where

\[
p(I) = \sum_v p(I|v) \tag{9.47}
\]

is the normalising denominator and \( p(v) \) is the prior model, which is a probabilistic description of the state of the deformable model to be estimated before any sensor data is collected. The sensor model \( p(I|v) \) expresses the stochastic processes that relate the original unknown state \( v \) to the input image data (sensor values) \( I \), in other words, the probability of producing the image \( I \) given a model \( v \). The posterior model \( p(v|I) \) describes the best estimate of \( v \) conditioned on the image data \( I \).

The initial snake serves as a prior model of the shapes and features of interest in the image. To cast the internal energy (9.2) of the deformable models into a prior distribution over expected shapes, a Boltzmann distribution [330] of the form

\[
p(v) = \frac{1}{Z} e^{-E_i(v)} \tag{9.48}
\]
ON DEFORMABLE MODELS

is employed with lower energy shapes being the more likely, where \( Z_i \) is a normalising constant (referred to as the partition function) and \( E_i(v) \) is the discretised version of \( E_{int} \) in equation (9.2). The sensor model may also be constructed on the basis of Boltzmann distribution as

\[
p(I|v) = \frac{1}{Z_p} e^{-E_p(v)}
\]

where \( E_p(v) \) is the discretised version of \( E_{ext} \) in equation (9.3). According to Bayes’ Theorem (9.46) the posterior distribution is obtained by combining (9.48) and (9.49)

\[
p(v|I) = \frac{p(I|v)p(v)}{p(I)} = \frac{1}{Z} e^{-E(v)}
\]

where

\[
E(v) = E_i(v) + E_p(v)
\]

is a discrete version of equation (9.8). Thus finding the minimum energy configuration of the snake is equivalent to finding the configuration of \( v \) which maximises the conditional probability \( p(v|I) \), which is known as the maximum a posteriori (MAP) estimate solution [330].

Although in many cases both the traditional model and the Bayesian models may be used to produce the same solution, the probabilistic framework has several advantages over the traditional model. Most notably, the probabilistic approach also provides a quantified measure of the uncertainty of the estimated shape parameters after the model is fitted to the image data, which may be used by higher level of processing at later stage [324].

It is also possible to extend this probabilistic framework by further assuming a time-varying prior model in conjunction with the sensor model, resulting in a Kalman filter [331]. Such a prior model describes the expected evolution of the deformable model shape parameters \( v \) over time. If the equations of motion of the dynamic deformable models are employed as the prior model, they become “Kalman snakes”, which are a promising method for tracking non-rigid objects in time-varying images [319].

9.7.2 Fourier Deformable Models

The traditional parametric snake model is a “global” boundary finding method in which the boundary is considered as a whole\(^2\). However, there is no global shape constraint or other \emph{a priori} information whatsoever built into the model. This sometimes makes it easily misled by spurious local structures in the image. Using elliptic Fourier descriptors as model parameters to represent closed and open contours, Staib and Duncan [327] proposed Fourier deformable models, in which both the boundary was considered as a whole and model-based global shape information was used.

\(^2\)Compared to common edge detectors such as Sobel and Prewitt edge operator, which glean edge information only from a small local neighbourhood of pixels, with high-order structure or organisation of the image completely ignored.
The elliptic Fourier representation of a closed contour \( \mathbf{v}(s) \) is given by

\[
\mathbf{v}(s) = \begin{bmatrix} x(s) \\ y(s) \end{bmatrix} = \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} a_k & b_k \\ c_k & d_k \end{bmatrix} \begin{bmatrix} \cos ks \\ \sin ks \end{bmatrix}
\] (9.52)

where \( s \in [0, 2\pi] \) is the independent parameter and

\[
a_0 = \frac{1}{2\pi} \int_0^{2\pi} x(s)ds \quad \quad b_0 = \frac{1}{2\pi} \int_0^{2\pi} y(s)ds
\]

determine the overall displacement of the contour and

\[
a_k = \frac{1}{\pi} \int_0^{2\pi} x(s) \cos ks ds \quad \quad b_k = \frac{1}{\pi} \int_0^{2\pi} x(s) \sin ks ds
\]

\[
c_k = \frac{1}{\pi} \int_0^{2\pi} y(s) \cos ks ds \quad \quad d_k = \frac{1}{\pi} \int_0^{2\pi} y(s) \sin ks ds
\]

represent ellipses in the parametric form. The geometric properties of each of the component ellipses, semi-major axis length, semi-minor axis length, rotation and phase shift, can be derived from \( a_k, b_k, c_k, d_k \). In other words, the contour can be geometrically interpreted as being composed of several ellipses, or viewed as a sum of rotating phasors, which individually define an ellipse and rotate with a speed proportional to their harmonic index.
This interpretation is demonstrated in Figure 9.3, where a contour is constructed from three different component ellipses. The dotted lines represent the phasors of the ellipses, which are combined to determine a point on the final contour at different times, denoted by $C_{31}, C_{32}, C_{33}$ respectively. The centre of the first ellipse $C_0$ is determined by $a_0, c_0$. $C_{ij}$ represents a point of the $i$th ellipse at time $j$, which also is the centre of ellipse at the next higher level, if any.

The Fourier coefficients $\{a_0, c_0, a_k, b_k, c_k, d_k\}$ form the parameter vector of the elliptic Fourier representation, $\Theta = (a_0, c_0, a_1, b_1, c_1, d_1, \cdots)$. These parameters follow a scale ordering, that is, low index parameters describe global shape properties and higher order parameters determines local deformations. In practice, the expansion series in equation (9.52) must be truncated. The truncation smooths the curve and limits the number of parameters, which determines the computational cost. But it will also decrease the accuracy of the representation. Choosing the number of harmonics is a trade-off between desired accuracy and degree of smoothing. Since most biological objects are relatively smooth, they are well represented by only a small number of harmonics. It was reported to range from 4 to 6 in [327].

Prior information is incorporated into the Fourier deformable model by associating the parameters with a probability distribution, which biases the model toward a particular (expected) range of shapes about a “mean” shape. The spread in the distributions accounts for the variability among the instances of the object shapes. A “mean” contour (solid line) with its variations (dotted and dashed lines) are shown in Figure 9.4. Here independence between the parameters and uniform distributions are assumed for the prior probability
of the parameters.

The prior probability distributions can be derived from the sample images of the object being delineated. The boundaries of the objects must first be identified manually. These boundaries are then decomposed into their model parameters and the statistics are collected. Prior probability distributions can then be estimated. When not enough sample images are available, uniform distributions can be assumed for the prior probability distributions.

Boundary finding is then formulated as an optimisation problem by fitting the model to the images in the parameter space and a Bayesian approach is then used to obtain the maximum a posteriori estimate of the boundary. Suppose $t_\Theta(x,y)$ represents an object shape template governed by a particular parameter vector $\Theta$. The goal here is to find which template $t_\Theta(x,y)$ the image $I(x,y)$ is most probably depicting, given the prior information and image data. In terms of probabilities, this is to evaluate $p(t_\Theta|I)$, the probability of the template given the image, and find the maximum over all $\Theta$. According to Bayes’ Theorem,

$$p(t_\Theta|I) = \frac{p(I|t_\Theta)p(t_\Theta)}{p(I)}$$

$p(I)$ is the prior probability of the image data and remains a constant for all $\Theta$. After eliminating $p(I)$ and taking the logarithm, the above equation reduces to

$$\ln p(t_\Theta|I) = \ln p(t_\Theta) + \ln p(I|t_\Theta)$$

The first term in equation (9.54) is the bias due to the prior probability. Assuming independence between the parameters, a multivariate Gaussian can be used for the prior probability of the $N$ parameters

$$p(\Theta) = \prod_{i=1}^{N} p(\theta_i) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(\theta_i-m_i)^2}{2\sigma_i^2}}$$

where $\theta_i$ represents the $i$th element of the parameter vector $\Theta$, $m_i$ is the mean of $\theta_i$ and $\sigma_i$ is the corresponding standard deviation. The second term in equation (9.54) is the contribution of the image information and can be derived as follows

$$p(I|t_\Theta) = \prod_{A} \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\frac{[I(x,y)-t_\Theta(x,y)]^2}{2\sigma_n^2}}$$

where $A$ denotes the whole image area and $\sigma_n$ is the variance of the noise $n$, which is defined as $n = I - t_\Theta$. It can be regarded as a correlation of the object shape template with the boundaries in the image and is thus a kind of matched filter [327, 332]. A continuous gradient ascent method [333] is then used to find the maximum of the posterior probability function, which is the best fit of the template to the image [327].

Fourier deformable models, as a boundary finding method, are aimed at situations
where there is some prior information about the global shape of the object but the information is not exact. They have been found to be relatively insensitive to the problems of broken boundaries and spurious edges from nearby objects. They are similar to the Hough transform in that they are based on template matching and find shapes by looking for a maximum in the parameter space. However, the advantage over the Hough transform is that the entire parameter space does not have to be constructed because of the use of local search in finding the maxima.

The flexibility of this model, both in terms of its probabilistic nature and the parametric representation, make it an attractive boundary finding method. However, this model has its own limitations. For example, each parameter has a small range within which the solution can be found to be reliable. Once the parameters vary beyond that range, the result will converge to false local minima corresponding to nearby features. This region of confidence about the true boundary depends on the quality of the image, the degree of smoothing, and the particular problem. Another limitation is that they are not suitable for describing some shapes, such as those with convolutions or sharp corners [334].

Fourier deformable models [327] and their extensions or variations are generally known as active shape models (ASMs) [334–338] (or statistical shape models). While traditional snakes control their shapes by imposing local smoothness constraints with a physical basis (in the form of internal elastic and bending force), ASMs accomplish this by constructing some prior models (point distribution models [334, 336, 337]), whose statistical properties are derived from a set of sample images, and then reducing the problem of boundary determination to finding the best possible “configurations” of the models under a maximum a posteriori (MAP) criterion using Bayes’ rule. Traditional snakes are so flexible that they are often trapped by local minima and converge to a suboptimal solution. On the contrary, however, ASMs are usually too “tight” to trace finer local details. Another limitation of the ASMs is that some landmark points usually need to be manually identified and annotated in the training images. Some of these limitations of the ASMs have been addressed by the recent development of hierarchical active shape models [289, 339, 340].

9.7.3 Gradient Vector Flow (GVF)

As mentioned in previous sections, traditional deformable models are not without limitations. Most notably, there are two problems that confront the practitioners of the traditional deformable models, that is, their sensitivity to initialisation and poor convergence to non-convex shapes.

The final contour extracted by a snake depends critically on the initial position and shape of the snake. In general, a snake must be placed close enough to the feature of interest otherwise the contour can be easily obstructed by unwanted features nearby. To address this problem, several methods have been proposed for increasing the capture range of the external force field, including the balloon model described in subsection 9.6.2 and distant attraction potential force of subsection 9.6.3. One problem with the balloon model
9.7. VARIATION AND EXTENSION OF THE ORIGINAL SNAKE MODEL

is that it must be prescribed to inflate or deflate depending on the prior information of the object of interest. Another problem is that the weighting parameter $\kappa_b$ must be selected with due care so that the inflation force is not too weak to pass by spurious edge points but not too strong to overwhelm legitimate edge segments. The distant attraction potential force greatly increases the capture range of a snake, but the second problem still remains unsolved, that is, the snake has poor performance in deforming towards regions with concave boundaries or indentations (such as the interior of a U-shape object).

To address the above two problems, Xu and Prince proposed the replacement of the external potential force of the original deformable models by a new external force they referred to as gradient vector flow (GVF) [175, 328] or generalised GVF [176, 317].

This new external force field is computed as a spatial diffusion of the gradient of an edge map derived from the image. This diffusion process causes the forces to propagate from strong edges not only to their immediate vicinity but also to a neighbourhood even far from the edges. This is like heat radiating from an energy source in the direction of maximum temperature decrease. Here the strong edge points can be thought as the “heat source”.

The gradient vector flow field is defined as a vector field $\mathbf{u}(x,y) = (u(x,y), v(x,y))$ that minimises the following functional

$$ I = \iint \mu |\nabla \mathbf{u}|^2 + |\nabla f|^2 |\mathbf{u} - \nabla f|^2 dxdy $$

(9.57)

where $\mu$ is a regularisation parameter which is dependent on the noise level in the image. A noisy image generally demands a large $\mu$. $f = f(x,y)$ is an edge map that is derived from the image $I(x,y)$, and usually defined like the external potential function in equations (9.4) to (9.7). A minimum of the functional (9.57) must satisfy the following Euler equation (Appendices A and B)

$$ \mu \nabla^2 \mathbf{u} - |\nabla f|^2 (\mathbf{u} - \nabla f) = 0 $$

(9.58)

which can be decoupled into the following two equations

$$ \mu \nabla^2 u - (f_x^2 + f_y^2)(u - f_x) = 0 $$

$$ \mu \nabla^2 v - (f_x^2 + f_y^2)(v - f_y) = 0 $$

(9.59)

where $\nabla^2$ represents the Laplacian operator. The physical intuition behind the gradient vector flow functional is now evident form the above two equations. As we can see, around strong boundaries, where the magnitude of the gradient is large, $u, v$ take on the values of $f_x, f_y$ respectively, $u = f_x, v = f_y$. This implies that $\nabla u = \nabla v = 0$. Therefore the two equations in (9.59) hold. For homogeneous areas, where $f_x = f_y = 0$, equation (9.59) reduces to Laplace’s equations, $\nabla^2 u = 0, \nabla^2 v = 0$. The gradient vector flow field in this area will be the result of “heat” flowing from the boundary of the region with the strong edge points regarded as the “heat source”. This spatial diffusion results in that
the external force not only exists around the strong boundaries but also extends far away from the boundaries into the homogeneous area.

To find a solution to equations (9.58), \( u \) is treated as time-varying and its partial derivative with respect to time \( t \), \( u_t \), and is then set to equal to the left hand side of equation (9.58), which becomes the following parabolic \(^3\) partial differential equation

\[
 u_t = \mu \nabla^2 u - |\nabla f|^2(u - \nabla f) \tag{9.60}
\]

These differential equations are referred to as generalised diffusion equations. Similar differential equations occur in physics applications such as heat conduction, fluid flow and neutron diffusion [175, 342]. The solution to the Euler equations (9.58) can then be obtained as the steady state solution of these parabolic differential equations, which are computed through a numerical iterative process using an explicit finite difference scheme as described in [175].

The GVF force field \( u \) obtained is then entered into the force balance equation (9.11) as a replacement of the traditional potential force (9.13). A snake under the influence of this new external force is referred to as a GVF snake. A GVF snake is then numerically simulated in the same way as the traditional parametric snake, governed by equation (9.20) except that the external force term \( F_{ext} \) is replaced with GVF force \( u \). A GVF snake has the desired properties that its capture range is greatly increased because of the diffusion process, which also allows the snake to slither toward concavities in the image it otherwise has difficulties moving into.

The generalised GVF snake model was proposed as an extension of the GVF snake model to further improve its ability to conform to long thin boundary indentations [176, 317]. A generalised GVF force field is defined as the equilibrium solution of the following partial differential equation

\[
 u_t = g(|\nabla f|)\nabla^2 u - h(|\nabla f|)(u - \nabla f) \tag{9.61}
\]

where \( g(|\nabla f|) \) and \( h(|\nabla f|) \) are two spatially varying weighting functions defined as

\[
g(|\nabla f|) = e^{-\frac{|\nabla f|}{K}} \tag{9.62}
\]

and

\[
h(|\nabla f|) = 1 - g(|\nabla f|) \tag{9.63}
\]

If we set

\[
g(|\nabla f|) = \mu \tag{9.64}
\]

---

\(^3\)Many practical problems encountered in science and engineering can be described as a differential equation of function \( u(x_1, x_2) \) of the form

\[
a u_{x_1 x_1} + 2b u_{x_1 x_2} + cu_{x_2 x_2} + du_{x_1} + ev_{x_2} + fu + g = 0,
\]

where \( a, b, c, d, e, f, g \) and \( g \) are given functions of \( x_1 \) and \( x_2 \). These partial differential equations are said to be: hyperbolic if \( ac - b^2 < 0 \); elliptic if \( ac - b^2 > 0 \); parabolic if \( ac - b^2 = 0 \) [341, 342].
and

\[ h(|\nabla f|) = |\nabla f|^2 \]  

(9.65)

it reduces to the original GVF snake model.

The GVF snake models solve most of the initialization and convergence problems for concave areas that usually hinder the traditional snake models. However, final performance of the snake still depends highly on good initialization of the snakes when the image is quite noisy or the magnitude of spurious edge segments are comparable to that of desired boundaries. Another downside of it is the high computational cost involved for the diffusion process.

9.8 Discussion

Deformable models are a set of promising and actively researched model-based image analysis techniques which combine physics, geometry and approximation theory. Snakes are planar deformable contours that move within an image and conform to its salient features under the influence of their inherent internal forces and external forces derived from the image. These forces are defined through the use of an energy minimisation framework.

Within the computer vision community and the like, low level visual tasks such as edge detection and motion tracking have long been regarded as autonomous bottom-up processes and many image processing techniques have been developed based on this recognition. Many problems have been encountered, however, in application of these model-free techniques to medical image analysis. For example, extracting boundary elements that belong to the same anatomic object can be quite challenging and reassembling these elements into a compact geometric representation is even more difficult due to the complexity and variability of the biological structures of interest, let alone noise and other artifacts causing the boundaries of structures to be indistinct and disconnected. Another disadvantage with these techniques is that decisions made at low level are irreversible resulting in mistakes being propagated to the subsequent higher level processes without the possibility of being rectified.

In contrast, deformable models integrate the bottom-up information derived from the image with top-down \textit{a priori} knowledge about the objects of interest in the same framework, yielding more reliable and consistent results. Two-dimensional, and more recently three-dimensional deformable models have found wide applications in a variety of medical image analysis problems, such as segmentation, visualisation, and tracking of biological structures such as the heart, brain, liver and so forth in images from different modalities [294].

The wide application of the deformable models indicates the important role that deformable models are currently playing and can potentially play in medical image analysis in the future. However, the traditional snake models have their own limitations, includ-
ing sensitivity to their initial conditions, limited capture range, and poor convergence to concave regions. Many variations and extensions have been proposed to address these problems, including balloon models, Fourier deformable models and GVF snake models, as described in the previous sections. Other notable variations or extensions include a dual active contour model [343,344] and B-spline snakes [345–349].

However, in some cases the initialisation problem remains where the images contain high level of noise and/or spurious local structures that are close to the desired boundaries. For these cases, the initial snake must be placed sufficiently close to the object of interest to provide good convergence. The initial contour is usually outlined manually by a user, which is a tedious and time-consuming task. In Chapter 10 we will present a contour extraction scheme based on snakes and the wavelet transform, in which snake initialisation is accomplished semi-automatically, which is less demanding of user interaction.

A snake placed in an noisy image may converge to a poor local minimum because the noise causes the existence of multiple minima in the snake energy minimising process. A scale-space method [174,350] was proposed to address this problem by first optimising a snake at a coarse scale with high level of smoothing to raise its attraction range and alleviate the interference from the noise, then tracking the optimal contour at finer scales with less smoothing to maintain boundary localisation accuracy. However, the amount of smoothing necessary for a noisy image is dependent on the noise level, which is usually unknown, and there is no established guideline on how to schedule the smoothing across scales. The wavelet transform provides a multi-scale representation of an image in the wavelet domain, in which a snake can be initialised and deform across different scales, from coarse to fine, to locate object boundaries of interest, as we can see in the contour extraction scheme based on snakes and wavelet transform in Chapter 10.

Another limitation with the traditional parametric deformable models is that they can only handle geometrically and topologically simple and consistent objects. Simultaneous treatment of multiple contour splitting and merging are not directly available in this framework. To address this problem an implicit model was proposed in [351], which provides topological and geometric flexibility through the use of level sets. More recently topologically adaptable snakes were proposed where a simplicial grid is superimposed over the image domain and used to iteratively reparameterise the deforming snakes so that they can be created and split into multiple parts or merge into other snakes depending on the image data [352,353].

The regularisation parameters are important to the performance of the traditional parametric deformable models and must be selected with due care. However, no “panacea” guidelines are available to determine the regularisation parameters. In practice they are estimated and fine-tuned on a trial and error basis.

In summary, medical image analysis has been both fostering and benefiting from the development of deformable shape models. With various medical image analysis techniques based on deformable models being applied in different imaging modalities, their usefulness
in routine clinical practice can be envisaged through extensive clinical validation and development and eventual deployment of clinically oriented systems for diagnostic radiology on the basis of these techniques can be expected in the foreseeable future.
Chapter 10

Contour Extraction Using Multiresolution Snakes

In this chapter a novel semi-automatic contour extraction scheme is proposed, which relies on the wavelet transform and snakes. It is motivated by the previous work done by other researchers in this area and the pilot studies in the previous chapters, which give us the perception that it is feasible to put the problem of extracting the contour of the prostate in a multi-scale framework provided by the wavelet transform.

Figure 10.1 illustrates an overview of the proposed contour extraction scheme, which consists of the following steps:

- First an ultrasound image is decomposed into edge maps at different resolutions using Mallat and Zhong’s discrete wavelet transform;
- Then some seed points are found in the coarsest edge map using both the modulus and phase angle of the wavelet transform;
- A snake initialised using these seed points evolves across the edge maps at different scales and finally converges to the contour of the prostate.

The wavelet transform provides a multi-resolution representation of an image. In particular, Mallat’s dyadic wavelet transform yields edge maps of an image at different scales. These edge maps can then be used as the external force fields for a snake, which deforms to extract the contour of the prostate.

It is important to initialise the snake fairly close to the desired features, and this will be demonstrated by the poor performance of a traditional snake, a balloon and a GVF snake. Unfortunately, there are no landmark features available in the ultrasound prostate images that could help automate the initialisation process. Therefore the snake is initialised using some “seed points” identified semi-automatically. To keep the interference of the noise to a minimum, seed point identification begins with the edge map at the coarsest scale. An anchor point is selected manually near the centre of the prostate region and the seed points are identified as the maxima along the radial profiles which emanate from the
10.1 The Motivation

As we have seen from Chapter 9, snakes have many important properties which make them a favourite methodology in applications such as image segmentation, motion analysis and tracking and so forth. Among these features, contour connectivity is particularly advantageous to the problem of segmentation, which is guaranteed by the inherent constraints due to the internal energy term of the snakes. If part of a snake finds a feature that could make its energy lower, the adjacent parts of the snake will be dragged towards this feature to find a possible continuation of it thanks to the internal forces, which penalise the formation of discontinuities. Just like the model based approaches described in the previous section, this overcomes the problems of edge linking and interpolation when edges are broken.

The external force plays an important role in driving the snake towards the desired features within the images, which are usually the locations of the maxima of the magnitude.

anchor point. After initialisation, the snake will evolve across the edge maps at different resolutions, from coarse to fine, and eventually extract the contour of the prostate.
of the gradients. The close relationship between the external force of the snake and zero-crossing identification in the Marr-Hildreth theory of edge detection can be established if the expression of the external force is revisited.

Suppose the external energy function of the snake is defined as in equation (9.7). Substitution of (9.7) into (9.13) gives the Gaussian potential force

\[ F_{\text{ext}} = \gamma \nabla |\nabla [G_\sigma(x,y) \ast I(x,y)]| \]  

(10.1)

and its magnitude is given by

\[ |F_{\text{ext}}| = \gamma |\nabla^2 [G_\sigma(x,y) \ast I(x,y)]| \]  

(10.2)

As we can see from the above equation, the magnitude of the external force takes its minima at the zero-crossings of \( \nabla^2 [G_\sigma(x,y) \ast I(x,y)] \), which is exactly what a LoG operator is expected to locate. The snake is driven towards these zero-crossings by the external force until it subsides. The relationship between snakes and zero-crossing identification in the Marr-Hildreth theory of edge detection is thus established.

This is illustrated in Figure 10.2. Figure 10.2(a) shows the pseudo-3D plot of an image which contains a vertical ramp edge obtained after smoothing a step edge with a Gaussian function. The external potential energy function of a snake is derived from the image as the negative magnitude of the gradient of the image, which is shown in Figure 10.2(b) as a pseudo-3D plot. Figure 10.2(c) shows the magnitude profiles of the external force, which is the gradient of the external potential energy function, thus equivalent to the Laplacian of the smoothed image. A central portion of the external force field is shown in Figure 10.2(d), where the external forces are presented as arrows whose length represent the magnitude of the external forces.

As we can see in Figure 10.2(b), the external energy possesses a valley, which is the location of the ramp edge. A snake placed on the slopes or near the ridge of the valley will slide into the valley to minimise its potential energy. This effect can also be explained in terms of the external forces, which push the snake toward the bottom of the valley, where the external forces reduce to zero, as shown in Figure 10.2(d).

Another fact that we can infer from Figure 10.2 is the effect of changing the standard deviation \( \sigma \) of the Gaussian function:

- Applying a Gaussian with a small \( \sigma \) results in relatively sharp edges, which allows a snake to follow the boundary accurately. On the other hand, the snake has to be placed fairly close to the edges because the external forces die out quickly, which means a limited “capture range” of the snake.

- On the contrary, a large value of \( \sigma \) creates a broad valley around the boundary and the external force of the snake is propagated far away from the edges. The “capture range” of the snake is thus extended. However, this extension of the “capture range”
10.1. THE MOTIVATION

Figure 10.2: The external force plays the important role in driving the snake towards the desired features within the images. (a) A step edge smoothed by a Gaussian function. (b) The external potential energy function. (c) Magnitude of the external potential force. (d) A central portion of the external force field.

does not come without a cost. Boundary localisation accuracy is undermined because nearby features are merged and displaced by the Gaussian function.

To get out of this dilemma, Kass et al. [174, 350] proposed a scale-space continuation method using Gaussian potential forces at different scales by varying $\sigma$. The snake can be attracted towards the desired boundaries from a long range by the Gaussian potential force at a coarse scale. The standard deviation $\sigma$ is then reduced to allow the snake converge to the desired boundaries at a finer scale. However, there are no established guidelines on how to select the value of $\sigma$ from one scale to another.

The previous research in this area and the pilot studies described in the previous chapters all suggest that the problem of extracting the contour of the prostate could be considered in a multi-scale context. The wavelet transform, especially the dyadic wavelet transform proposed by Mallat and Zhong [195] provides exactly such a sound mathematical foundation, via which the multi-resolution representation of the original image at dyadic scales can be obtained and employed later in the contour extraction scheme.
10.2 Multi-Scale Edge Detection

Mallat and Zhong’s dyadic wavelet transform \[195\] has several properties which make it a perfect choice to fit in with our contour extraction scheme. The smoothing function\(^1\) \(\vartheta(x, y)\) associated with the dyadic wavelet transform is an approximation to a Gaussian function. Since the quadratic spline wavelet is the first order derivative of the smoothing function \(\vartheta(x, y)\), the components of the wavelet transform of an image are actually in proportion to the components of the gradient of the image smoothed by \(\vartheta(x, y)\), as we demonstrate below. This means the dyadic wavelet transform closely resembles gradient edge detectors in the way they locate edge points as maxima of the magnitude of the gradient. Moreover, edge maps of different resolutions are presented from coarse to fine at dyadic scales in a systematic way guaranteed by the dyadic wavelet transform framework.

The dyadic wavelet transform of an image \(I(x, y) \in L^2(\mathbb{R}^2)\) at scale \(2^j\) has two components, the horizontal component denoted by \(d_{2j}^H(x, y)\) and the vertical component \(d_{2j}^V(x, y)\). They are defined by the convolution with the two directional (horizontal and vertical) wavelets at scale \(2^j\), \(\psi_{2j}^H(x, y)\) and \(\psi_{2j}^V(x, y)\) respectively

\[
\begin{align*}
  d_{2j}^H(x, y) &= \psi_{2j}^H(x, y) * I(x, y) \\
  d_{2j}^V(x, y) &= \psi_{2j}^V(x, y) * I(x, y).
\end{align*}
\]

(10.3)

The two directional wavelets at scale \(2^j\) are dilated versions of their corresponding mother wavelets

\[
\begin{align*}
  \psi_{2j}^H(x, y) &= \frac{1}{2^j} \psi^H \left( \frac{x}{2^j}, \frac{y}{2^j} \right) \\
  \psi_{2j}^V(x, y) &= \frac{1}{2^j} \psi^V \left( \frac{x}{2^j}, \frac{y}{2^j} \right),
\end{align*}
\]

(10.4)

which are given by the partial derivatives of the associated spline smoothing function \(\vartheta(x, y)\) as follows

\[
\begin{align*}
  \psi^H(x, y) &= \frac{\partial \vartheta(x, y)}{\partial x} \\
  \psi^V(x, y) &= \frac{\partial \vartheta(x, y)}{\partial y}.
\end{align*}
\]

(10.5)

If we denote \(\vartheta_{2j}(x, y) = \frac{1}{2^j} \vartheta \left( \frac{x}{2^j}, \frac{y}{2^j} \right)\) Equation (10.3) can be rewritten as

\[
\begin{align*}
  \begin{pmatrix}
    d_{2j}^H(x, y) \\
    d_{2j}^V(x, y)
  \end{pmatrix} &= \begin{pmatrix}
    \psi_{2j}^H(x, y) * I(x, y) \\
    \psi_{2j}^V(x, y) * I(x, y)
  \end{pmatrix} \\
  &= 2^j \begin{pmatrix}
    \frac{\partial}{\partial x} (\vartheta_{2j} * I)(x, y) \\
    \frac{\partial}{\partial y} (\vartheta_{2j} * I)(x, y)
  \end{pmatrix} \\
  &= 2^j \nabla (\vartheta_{2j} * I)(x, y).
\end{align*}
\]

(10.6)

\(^1\)Definition of the wavelets and the associated smoothing function are given in Appendix A of [195].
10.2. MULTI-SCALE EDGE DETECTION

As we can see in (10.6), the two components of the dyadic wavelet transform of an image $I(x, y)$ at scale $2^j$ are proportional to the components of the gradient of a blurred version of the original image, which is smoothed by the corresponding scaling function $\vartheta_{2^j}$.

Thus the modulus of the wavelet transform is given by

$$M_{2^j}(x, y) = \sqrt{|d_H^{2^j}(x, y)|^2 + |d_V^{2^j}(x, y)|^2} \quad (10.7)$$

and the phase angles of the wavelet transform vectors are

$$A_{2^j}(x, y) = \arctan\left(\frac{d_V^{2^j}(x, y)}{d_H^{2^j}(x, y)}\right) \quad (10.8)$$

Shown in Figure 10.3 is one of the ultrasound images of the prostate “us1” decomposed up to scale $2^4$ via the dyadic wavelet transform. Figure 10.3(a)-(d) are the magnitudes of the transform, from $M_{2^1}(x, y)$ up to $M_{2^4}(x, y)$. Figure 10.3(e)-(h) represent the corresponding angles. As we can see in Figure 10.3(a)-(d), the notable gray scale transitions within the original image, which can be attributed to the salient features such as the prostate gland boundaries, are present as local maxima in the modulus of the wavelet transform, which are identified as the bright pixels. The transform thus provides us with the edge maps of the original image at different resolutions from the finest to the coarsest. This is one of the most attractive characteristics of this transform.

Note how the contours of the prostate are well pronounced in the coarser (large scale) edge maps, such as that of scale $2^3$ and $2^4$. The other side of the coin is that they are also blurred, expanded and displaced due to the smoothing of the scaling functions. In contrast, at small scales, the contours of the prostate are well localised. However, the noise also manifests itself in a random pattern in these edge maps, making it difficult to isolate “legitimate” edge points from the spurious ones.

This observation suggests that a rough boundary of the prostate could be obtained from the edge map at the coarsest level to avoid the interference of the noise. Then the boundary can be refined by incorporating the information from the edge maps at smaller scales. This approach is comparable to Kass’ scale-space continuation method [174, 350]. Similar ideas to this scale by scale paradigm can also be found in Rosenfeld and Vanderberg’s coarse-fine template matching [213].

The phase angle of the two components of the wavelet transform also carries important information, which gives the orientation of the gradient vector at each pixel. The information of both the modulus and the angle will be employed later to initialise a snake, which is expected to deform across the edge maps at different scales and eventually converge to the contour of the prostate. This will be discussed next.
Figure 10.3: The dyadic wavelet transform of the prostate ultrasound image “us1”. (a)-(d) Modulus $M_{2^j}(x, y)$, $j = 1, 2, 3, 4$. (e)-(h) Angle $A_{2^j}(x, y)$, $j = 1, 2, 3, 4$. 
10.3 Snake Initialisation: The Importance

Applying the dyadic wavelet transform to an image is comparable to taking the gradient after smoothing it by a Gaussian function. Thus it is sensible to designate the modulus of the dyadic wavelet transform as the external potential energy function of a snake, that is,

\[ P(x, y) = -\gamma M_{21}(x, y) \]  \hspace{1cm} (10.9)

where \( \gamma \) is a positive weighting parameter. This is equivalent to using the Gaussian potential energy function as defined by equation (9.7).

Traditionally, the external potential function of a snake is defined only once and for all at a fixed resolution\(^2\) and the snake evolves only within this specific external force field. In other words, the external potential function, once decided on, cannot be changed during the deformation of the snake, making the snake either settle down on rather coarse and displaced boundaries when the Gaussian smoothing is strong, or compromised by the interference of the noise otherwise. In contrast, it is feasible to get out of this dilemma by switching the external potential function of a snake between the edge maps at different resolutions obtained from the dyadic wavelet transform, and let it deform across different scales, as implied in equation (10.9).

The deformation of such a multi-scale snake, whose external potential function is “switchable”, begins with the coarsest edge map. Once the snake settles down within this external force field, the change of the external potential function to the next finer scale will perturb the equilibrium formerly reached. Therefore the snake will restart its deformation and reach a new equilibrium and result in a contour with finer details. This process repeats until it arrives at the finest edge map. The migration of the snake from coarse to fine scales is thus carried out in a systematic way provided by the dyadic wavelet transform.

Since the boundaries are displaced and merged in the coarsest edge map, the snake must be initialised reasonably close to the desired boundary. Otherwise the snake might deviate significantly from the “ground-truth” contour. To demonstrate the importance of initialisation for the snakes, a traditional snake, a balloon, and a GVF snake are applied to image “us1” and their deformation are examined next. The parameters of the snakes are summarised in Table 10.1.

10.3.1 Traditional Snake

As the first example the deformation of a traditional snake is demonstrated in Figure 10.4. The regularisation parameters of the snake are \( \alpha = 0.1, \beta = 0.1, \gamma = 2.0 \) respectively. The initial contour of the snake is defined to be a circle and placed in the dyadic wavelet transform domain of the original image “us1”. The modulus of the wavelet transform\(^2\) Determined by a specific \( \sigma \) value in the Gaussian potential function, for example.
CONTOUR EXTRACTION USING MULTIRESOLUTION SNAKES

<table>
<thead>
<tr>
<th>SNakes</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>$\kappa_b$</th>
<th>$\mu$</th>
<th>Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>0.1</td>
<td>0.1</td>
<td>2.0</td>
<td>-</td>
<td>-</td>
<td>200</td>
</tr>
<tr>
<td>Balloon</td>
<td>0.1</td>
<td>0.1</td>
<td>2.0</td>
<td>0.1</td>
<td>-</td>
<td>200</td>
</tr>
<tr>
<td>GVF</td>
<td>0.1</td>
<td>0.1</td>
<td>2.0</td>
<td>0.2</td>
<td>-</td>
<td>200(^\dagger)</td>
</tr>
</tbody>
</table>

\(^\dagger\)Extra iterations for computation of the GVF field.

Table 10.1: Parameters of a traditional snake, a balloon and a GVF snake that were applied to image “us1” as an example to demonstrate the importance of initialisation of the snakes.

at scale $2^4$ is designated as the external potential function of the snake, as shown in Figure 10.4(a).

The deformation process of the snake is shown in Figure 10.4(b) and the final shape of the snake after 200 iterations is shown in Figure 10.4(c). As we can see, the snake has no difficulties in deforming into the anterior and posterior of the prostate contour. However, it is not the case when converging towards the two lateral lobes. This can be attributed to the fact that the upper and lower portion of the circle are fairly close to the prostate contour, whereas its left and right portion are far away from it.

This effect can also be explained after examining Figure 10.5, which shows the external force field. The upper portion of the snake is quickly pushed towards the anterior of the prostate contour by the strong external forces present. A similar situation prevails at the posterior of the prostate contour. On the other hand, the strong forces around the hyperechoic features within the prostate trap the snake and impede its further deformation into the lateral lobes of the prostate.

The poor performance of the traditional snake in this case naturally leads one to think of the variants of the deformable models, such as balloons [303, 304, 313, 320] or the GVF snake [175, 328]. The typical results of applying a balloon and a GVF snake are shown in Figure 10.6 and Figure 10.7 respectively.

10.3.2 Balloon

Figure 10.6(a) shows the deformation process of a typical balloon with snake parameters $\alpha = 0.1, \beta = 0.1, \gamma = 2.0$ and inflation force parameter $\kappa_b = 0.1$. The final shape of the balloon after 200 iterations is given in Figure 10.6(b).

As we can see, the performance of the balloon is slightly better than that of the snake in Figure 10.4 in terms of approaching the lateral lobes of the prostate contour. However, this marginally better convergence does not come without a cost. Although the balloon has managed to move pass the two hyperechoic regions, its entanglement has not completely disappeared, and is manifested as the loops or crossovers formed nearby, which are not desirable at all. Moreover, overshoots at the top left portion and the posterior are also not desirable, which can be explained by the fact that the inflation force overwhelms the internal force plus the external force there.
Figure 10.4: A snake initialised not sufficiently close to the desired boundaries has poor convergence. (a) A snake initialised as a circle at scale $2^4$ in the wavelet transform domain of image “us1”. $\alpha = 0.1, \beta = 0.1, \gamma = 2.0$. (b) The snake has no difficulties in deforming into the anterior and posterior of the prostate contour. However, it is not the case when moving towards the two lateral lobes. (c) Final form of the snake after 200 iterations.
10.3.3 GVF Snake

The result of applying a typical GVF snake is not encouraging either, as we can see in Figure 10.7(b), where a GVF snake is shown after 200 iterations. The deformation detail of the GVF snake is shown in Figure 10.7(a), whose parameters are still $\alpha = 0.1, \beta = 0.1, \gamma = 2.0$. The GVF regularisation parameter $\mu$ is set to 0.2, and the number of iterations to compute the GVF field is 50.

As we can see in Figure 10.7(a), it only takes the first few iterations for the GVF snake to arrive at a form which is not a far cry from its final form, indicating that a GVF snake converges faster than a traditional snake or a balloon. However, its final performance is no better than that of the snake in Figure 10.4. Moreover, the “saw-teeth” along the final contour are an undesirable artifact attributable to the time-consuming GVF computation process.

10.3.4 Discussion

The balloons and GVF snakes are expected to approach the desired object from a fairly long distance. However, this is attainable only if the image is relatively “noise-free” or “interference-free”. For this case, since the magnitude of the two hyperechoic regions within the prostate is comparable to that of the prostate boundary (even stronger than some parts of it), the disappointing performance of the balloons and GVF snakes is thus
Figure 10.6: The performance of a balloon is no better than that of a traditional snake in this case. (a) Typical deformation process of a balloon initialised as a circle at scale $2^4$ in the wavelet transform domain of image “us1”. $\alpha = 0.1, \beta = 0.1, \gamma = 2.0, \kappa_b = 0.1$. (b) Final form of the balloon after 200 iterations.

not totally unexpected.

All the results from the trials described above suggest that a snake be initialised reasonably close to the prostate boundary. The problem here is how to determine that a snake is sufficiently close to the contour of the prostate, which is exactly what the snake is expected to locate. In the ultrasound prostate images there is no landmark whatsoever available, which could be used as a starting point or reference point in determining a snake’s initial position. So a semi-automatic initialisation method is adopted here, in which both the modulus and the angle of the wavelet transform will be employed to initialise a snake, as described in the next section.
Figure 10.7: The performance of GVF snakes are not encouraging either. (a) Typical deformation process of GVF snake. $\alpha = 0.1, \beta = 0.1, \gamma = 2.0$. GVF regularisation parameter $\mu = 0.2$, and the number of iterations computing the GVF is 50. (b) Final form of the GVF snake after 200 iterations.

10.4 Snake Initialisation: A Semi-Automatic Centre-Based Approach

To initialise a snake reasonably close to the prostate boundary, some preliminary boundary points (referred to as “seed points” in the following text) are first identified in the coarsest edge map. In the following a semi-automatic centre-based approach is adopted in locating these seed points. Similar centre-based strategies can be found in other image segmentation practices where the objects of interest resemble an oval or other similar shapes [264, 266, 267, 277–281].
10.4. SNAKE INITIATION: A SEMI-AUTOMATIC CENTRE-BASED APPROACH

10.4.1 The Radial Profiles

First, the centre of the prostate region is roughly specified by the user. This is referred to as the “anchor point” in the following text. The anchor point is the origin of a set of radial lines equally spaced in angle, as shown in Figure 10.8(a). The radial profiles are extracted as gray scale value of the pixels along the radial lines. The user also roughly designates the first seed point on the boundary of the prostate, which is denoted as a cross in Figure 10.8(a). Typical radial profiles are shown in Figure 10.8(b), from top to bottom, corresponding to the radial lines in Figure 10.8(a), rotating clockwise from the first line marked with the cross.

As we can see in Figure 10.8(b) the prostate boundary points are present as the local maxima along the radial profiles. Moreover, the loci of the prostate contour are vaguely discernible if we fix our eyes on the maxima on each line and “link” them subjectively from top to bottom to form an imaginary curve. In other words, proceeding radially from the anchor point, a seed point can be identified by searching for a local maximum, which is larger than its two closest neighbours and strictly larger than at least one of them.

However, a further examination of Figure 10.8(b) reveals a problem. As we can see, some of local maxima corresponding to the desired seed points are of large magnitude, but others are very weak in strength, especially around the area where the prostate boundary is rather vague. If these maxima were the only maximum along each radial profile, this would not be a problem. Unfortunately, usually there are other local maxima along each radial profile, and sometimes the magnitude of the local maxima corresponding to the would-be seed points is rather weak compared to that of the spurious features the radial lines traverse. For example, the boundary points around the apex of left lateral lobe (at the lower left quadrant of the image) have no advantage in strength over the points close to the anchor point, which apparently are not part of the prostate boundary.

Therefore, given the fact that the local maxima corresponding to the desired seed points have no apparent advantage in magnitude over the “noise”, it is not a sensible approach to search “globally” along the radial profiles for the boundary points, as in many centre-based strategies [264, 266, 267, 277–281], since this will not only significantly increase the searching overhead in total, but also make the searching procedure subject to the interference from the spurious features along the radial lines. Ideally the seed points should be sought within a small neighbourhood to fend off interference and increase efficiency. How can this neighbourhood be adaptively defined?

10.4.2 The Adaptive Neighbourhood

As we can see in Figure 10.8(b), every two maxima corresponding to the adjacent desired seed points are fairly close to each other along the imaginary curve (loci of the prostate contour) because of the fact that the contour is relatively continuous. Therefore the next seed point should be sought somewhere close to the previous one.

It is a known fact that an edge point takes the maximum value in the magnitude of
Figure 10.8: The centre based approach: extracting the radial profiles. (a) Radial lines emanating from an anchor point specified by user. (b) The radial profiles extracted, from top to bottom, corresponding to the radial lines in (a) rotating clockwise from the first line marked with the cross in (a). Note the prostate boundary points present themselves as the local maxima along the radial profiles.

The gradient, which points to the direction where the largest gray scale transition occurs. This is illustrated in Figure 10.9, where dotted lines represent the boundary formed by two relatively iso-intensity areas. As we can see, at each edge point, the edge vectors are tangent to the iso-intensity contours, and perpendicular in either direction to the gradients and roughly point to the position of the adjacent edge points.
10.4. SNAKE INITIALISATION: A SEMI-AUTOMATIC CENTRE-BASED APPROACH

Figure 10.9: Gradients and edge vectors are perpendicular to one another. An edge point is identified as a point which takes the maximum value in the magnitude of the gradients. The edge vectors are perpendicular to the gradients and roughly point towards the adjacent edge points.

In the case of the dyadic wavelet transform, the edge vectors are perpendicular to the wavelet transform vectors (gradients or external forces for the snakes), that is,

\[ \theta_{2j}^e(x, y) = A_{2j}(x, y) \pm \frac{\pi}{2} \]  

(10.10)

The edge vectors of image “us1” at scale \(2^4\) are illustrated in Figure 10.10. Note how the edge vectors are perpendicular to the corresponding external forces (gradients) in Figure 10.5. Moreover, as expected, the edge vector of each seed point roughly points to the location of the next desired seed point.

To adaptively define the neighbourhood within which the seed points will be searched for, two lines are drawn out from the current seed point to the next radial line, as shown in Figure 10.11(a). One of them is extended along the direction of the edge vector. The other one is the perpendicular line taken from the current seed point to the next radial line. The intersection points of these two lines with the next radial line forms a neighbourhood within which the next would-be seed point will “most likely” be found.

Figure 10.11(b) illustrates an ideal scenario, where the solid line represents the radial profile and the two vertical dotted lines define the “most likely” neighbourhood within which the desired seed point will be searched for. Note how the local maximum in the middle, which corresponds to the desired seed point, compares unfavourably in magnitude to other spurious maxima. If it were searched “globally” along the whole radial profile,
the seed point would not be correctly identified.

However, searching in the “most likely” neighbourhood only is not without its own problem. For the particular case shown in Figure 10.11(b), fortunately the desired seed point lies right within the “most likely” neighbourhood. As said before, however, this is an ideal case only. It is not rare for these desired seed points to fall outside this neighbourhood, as we can see in Figure 10.11(a). Therefore, it is necessary to find a strategy which somehow favours these points and finally claims them as seed points successfully.

Figure 10.11(b) illustrates a sensible approach to this problem. The points outside the “most likely” neighbourhood are adversely weighted by a Gaussian function (dotted line) to suppress the possible interference of the spurious local maxima. The farther away a point from the neighbourhood, the stronger the suppression. This way the local maxima far away from the neighbourhood, which are most likely to be spurious ones, will not challenge the local maxima within or close to it, which are most likely to be the true seed points. Even if some desired seed points fall outside the neighbourhood, they still can be successfully asserted as long as they are close enough to the neighbourhood so that the adverse weighting hardly affects their magnitude. The radial profile after adverse weighting is shown as a stem plot in Figure 10.11(b). As we can see, the maximum in this weighted profile is then successfully found as the next seed point.
10.4. SNAKE INITIALIZATION: A SEMI-AUTOMATIC CENTRE-BASED APPROACH

10.4.3 Locating the Seed Points

Figure 10.12(a) demonstrates a typical running of the seed point finding strategy at scale $2^4$ of the modulus of the dyadic wavelet transform. The plus sign indicates the anchor point and the cross is the first seed point, both selected by the user. User interaction is kept to a minimum by only requiring that these two points be specified by user. After

Figure 10.11: Locating the seed points by searching for local maxima in radial profiles. (a) The edge vectors and the perpendicualrs. (b) Searching for the local maximum.
the anchor point and the first boundary point are specified the seed points will be found automatically in chains with the radial lines rotating clockwise.

Figure 10.12(b) shows all the seed points successfully identified applying this strategy. Note how the first seed point finally identified deviates a little from its rough initial position (shown as the cross) designated by the user. As said before, it can be selected rather casually by the user, as long as it is not unreasonably far away from the true boundary of the prostate. The seed point finding strategy will ultimately locate it as the local maximum along the radial profile.

In summary, a snake must be initialised reasonably close to the boundary of the
prostate in order for the snake to eventually delineate it. Some preliminary seed points are used to initialise the snake, and a semi-automatic centre-based strategy is adopted to locate them. After an anchor point and a boundary point are specified by the user, the seed points will be identified automatically along the radial lines emanating from the anchor point and rotating clockwise.

Both the modulus and the angle of the dyadic wavelet transform are employed to locate the seed points. The seed points are identified as local maxima along the radial profiles in the coarsest edge map, that is, the modulus of the dyadic wavelet transform. The information on the angle of the wavelet transform is used to define a “most likely” neighbourhood around which the next desired seed point will be sought. After being initialised with these seed points, the snake evolves across the edge maps at different resolutions and finally stabilises and rests on the contour of the prostate, as shown in the next section.

### 10.5 Multi-Scale Snakes: Extracting the Prostate Contour

It is essential to the contour extraction scheme that a snake be initialised by some seed points, which are identified semi-automatically by a centre-based approach, as discussed in the previous section. Even more important, however, is the conception of switching the external force field of the snake between the edge maps at different resolutions, that is, making it deform across different scales of the dyadic wavelet transform, as indicated in equation (10.9).

Such a multi-scale snake, whose external force field can be switched, distinguishes itself from a traditional snake, whose external potential function is defined only once and for all at a fixed resolution, which results in the dilemma that extending capture range and keeping boundary localisation accuracy cannot be achieved at the same time.

After the seed points are identified, a snake initialised with them will deform and finally reach an equilibrium within the current external force field, which is at the coarsest scale. Since at this scale the boundaries are usually rather smoothed, expanded and displaced, the contour extracted by the snake at this stage is not necessarily accurate. To further “refine” the contour, the external force field of the snake is switched to the next finer scale. Thus the equilibrium formerly reached will be perturbed and the snake will restart its deformation and finally settle down and produce a contour with finer details.

To demonstrate the contour extraction scheme which incorporates the conception of multi-scale snakes it is applied to image “us1” as an example. The seed points identified

<table>
<thead>
<tr>
<th>SCALES</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\gamma$</th>
<th>ITERATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s = 2^s$</td>
<td>0.1</td>
<td>0.1</td>
<td>2.0</td>
<td>100</td>
</tr>
<tr>
<td>$s \leq 2^4$</td>
<td>0.2</td>
<td>0.2</td>
<td>1.0</td>
<td>100</td>
</tr>
</tbody>
</table>

**Table 10.2**: Parameters of a multi-scale snake that is applied to image “us1”.
CONTOUR EXTRACTION USING MULTiresolution SNAKES

Figure 10.13: Multi-scale snake deforming and extracting the contour of the prostate in image “us1”. (a) Snake evolving at scale $2^4$. (b) Zoom-in of part of the snake at scale $2^4$. (c) Snake deforming at scale $2^3$. 
10.5. **MULTI-SCALE SNAKES: EXTRACTING THE PROSTATE CONTOUR**

Figure 10.14: Final contour extracted for image “us1”.

at scale $2^4$, which are shown in Figure 10.12, are used to initialise the snake, which starts evolving from the edge map at that scale (Figure 10.13(a)). The control parameters of the snake are summarised in Table 10.2.

Figure 10.13(a) shows the deformation process of the snake at scale $2^4$. As we can see, the snake first contracts and smooths itself due to the influence of its internal energy. Gradually it is attracted to the prominent edges within the image due to the external potential force. This is more evident in Figure 10.13(b), in which the zoom-in of the lower right part of the snake is shown. The bright line represents the initial position of the snake. As we can see, the internal forces make the snake contract and move to the left of its initial position. After a few iterations, the external force gradually takes the reins and drives the snake towards its initial position. The snake keeps moving towards the salient edges and finally stabilises.

Once the snake reaches equilibrium at scale $2^4$, its external force field will be “switched” to the edge map at scale $2^3$ (Figure 10.3(c)). The equilibrium will be perturbed and the snake starts evolving within the edge map at scale $2^3$, as shown in Figure 10.13(c). Note how the snake is refined and directed towards finer features along the prostate contour. Although the deformation of the snake after each iteration seems trivial, as indicated by the merged curves, there is still noticeable deviation between its final form and original one.

The final form of the snake is shown in Figure 10.14. As we can see, the contour extracted by the snake closely matches the prostate boundary discernible in the original image.

Figure 10.15 shows another example of application of the proposed contour extraction scheme. The seed points identified at scale $2^4$ of the dyadic wavelet transform of image “vs01” are shown in Figure 10.15(a). Note the “overshoot” seed point identified around
CONTOUR EXTRACTION USING MULTIRESOLUTION SNAKES

Figure 10.15: Results of applying the proposed contour extraction scheme to image “vs01”. (a) Seed points identified at scale $2^4$. (b) Snake deformation at scale $2^4$. (c) Snake deformation at scale $2^3$. [Images of the results shown in the figure]
10.6 DISCUSSION

Figure 10.16: Final contour extracted for image “vs01”.

the left lateral lobe which deviates from the “ground-truth” boundary due to the broken edge segment there. Snake deformation at scale $2^4$ is shown in Figure 10.15(b). Note how the snake rectifies the “overshoot” seed point thanks to its internal forces, which penalise abrupt changes on the curve. Figure 10.15(c) shows the snake evolving at scale $2^3$.

The final extracted contour is superimposed on the original image as a bright curve, shown in Figure 10.16. As we can see, the contour extraction scheme successfully defines the boundary of the prostate, which coincides with the subjective contour a human observer would infer from the image.

10.6 Discussion

In this section some issues specific to the semi-automatic contour extraction scheme will be discussed.

10.6.1 Decomposition Scale of the Dyadic Wavelet Transform

The maximum decomposition scale of the wavelet transform is determined by the image size. However, at the maximum scale, the image becomes a feature-free blank image whose pixel value is the average grey scale of all the pixels of the original image. Therefore, the images never need to be decomposed to the maximum scale available. Since the higher the scale, the greater the extent of smoothing caused by the scaling function, smaller features will be displaced and merged. Thus the real maximum decomposition scale is also related to the size of features of interest, in this case, the prostate. In practice, the maximum decomposition scale of the dyadic wavelet transform is found to be

$$s_{\text{max}} = 2^{\lceil \log_2 \min(M,N) \rceil - 4}$$

(10.11)
where \( \lceil x \rceil \) denotes ceil of \( x \), that is, the smallest integral value not less than \( x \). \( M \) and \( N \) are the number of rows and columns of the image. For the two images taken as example in the previous section, the real maximum decomposition scale is \( 2^4 \).

10.6.2 Control Parameters of the Snake

At the highest scale (\( 2^4 \) for the two examples in the previous section), \( \alpha \) and \( \beta \) are set to 0.1 and \( \gamma \) is set to 2.0, as listed in Table 10.2. This is to emphasize the image force at higher scales. Since the snake is more stretchy and less rigid and most of the noise has been removed at this scale (compared to the snake at the next stage), the snake will be driven to the rough contour of the prostate as quickly as possible.

At lower scales \( \alpha \) and \( \beta \) are increased to 0.2 to emphasize the “tension” and “rigidity” of the snake and \( \gamma \) is reduced to 1.0. This makes the snake resistant to the interference from the noise or other spurious features, which are more prevalent than at higher scales.

10.6.3 The Anchor Point

It is essential to the contour extraction scheme that the anchor point be specified at a sensible position. Since a centre-based approach is adopted to locate the seed points and the radial lines are equal spaced in angle, the anchor point should be roughly around the centroid of the prostate region. Otherwise at the parts that are far from the anchor point, the adaptive neighbourhood will be larger, making it more likely for the seed point locating strategy to go astray.

Can the anchor point (and the first seed point) somehow be found automatically so that our contour extraction scheme becomes fully automatic? Unfortunately, the answer is no. There is no landmark point or feature whatsoever in the ultrasound images of the prostate that can be used to locate the centre of the prostate region. To determine the centre of the prostate, the boundary of the prostate need to be defined, which is exactly what the scheme is expected to extract. This makes a classic “egg and chicken” problem.

On the other hand, however, it is pointless to design a contour extraction scheme which involves extensive interaction. By requesting only two points (the anchor point and the first seed point) be specified manually, user interaction is kept to a minimum and the rest of the seed points are found automatically.

10.6.4 The Seed Points

Can the seed points be used to directly form the contour of the prostate, and thus skip the procedure of snake deformation? After all, most of the seed points are around the prostate boundary. As we can see in the previous section, the deviation of the contour formed by the seed points (that is, the initial snake contour) and the final contour extracted by the snakes is not negligible. Moreover, “overshoot” seed points like that in Figure 10.15 are not uncommon. Obviously the final contour extracted by the snake is more accurate.
On the other hand, this implies that the seed points do not need to be very accurately located. In other words, localisation of the seed points and hence initialisation of the snake are “fault-tolerant”. If part of a snake goes astray initially, the adjacent parts of the snake will help rectify this situation due to effect of the internal force of the snake and eventually the contour of the prostate can still be accurately located.

10.7 Final Remarks

In Chapter 6 a speckle reduction method based on Mallat and Zhong’s dyadic wavelet transform was presented, which fits in with the general “picture” of computer aided analysis of medical images as a preprocessing procedure (see Figure 1.1). Intuitively, speckle reduced ultrasound images would be used as a starting point for the subsequent segmentation process. Unexpectedly, the contour extraction method presented in this chapter operates on the original “noisy” images, which seems surprising on first thought.

Since our contour extraction method also exploits the dyadic wavelet transform, whose associated smoothing function closely approximates a Gaussian function (see section 10.2), decomposition of an image via this transform inherently involves a denoising process, which is comparable to Gaussian filtering. Therefore, the resultant edge maps at different resolutions are relatively noise-free, especially at the coarsest level, which is then assigned as the initial external force field of the snake. In summary, thanks to the “built-in” noise reduction capability of the dyadic wavelet transform, our contour extraction scheme can directly operate on the original speckled images, which is not unreasonable.

However, this is not to say that the purpose of our noise reduction method in this case is simply visualisation (although admittedly visualisation is still a crucial part of a physician’s clinical practice for medical image analysis, as it is unlikely that automated segmentation methods will completely replace physicians in the foreseeable future). Apart from the proposed contour extraction scheme, which is specific to ultrasound images of the prostate, there are a variety of segmentation methods for medical images (see [291]) which can benefit from the preprocessing procedure of noise reduction. As we have demonstrated in Figure 6.16, it is more effective to apply any histogram-based or other pixel-based segmentation methods to the speckle-reduced image than to the original noisy image.

Although our study has been focused on ultrasound images of the prostate, the proposed methodologies, namely the noise reduction scheme and contour extraction method based on dyadic wavelet transform and snakes, may remain valid and useful for filtering or segmentation of ultrasound images in many other applications or medical images from other imaging modalities. While the wavelet and snake methodology is generic, the radial-profile-based snake initialisation approach is specific to objects of simple shape like the prostate. It has limitations when dealing with objects with concave structures, which may land the radial search strategy into trouble.

The notable difference between our contour extraction scheme and other multiresolu-
tion segmentation approaches based on snakes [283, 284, 354–356] lies in how the wavelet transform is exploited. In our contour extraction method the dyadic wavelet transform provides multiresolution edge maps, which are used as the external force field of a snake. Whereas focus is placed on parameterising the snake contour based on a one-dimensional dyadic wavelet transform involving Coiflet wavelets in [354–356] or biorthogonal spline wavelets with 1 vanishing moment (decomposition) and 3 vanishing moments (reconstruction) in [283, 284]. In these cases, an extra inflation (balloon) force is implicitly [354–356] or explicitly [283, 284] introduced, which makes the snake more vulnerable around weak features and prone to suffer overshoots (see also subsection 10.3.2), especially in [354–356], where the snake shape is unrestricted.

That said, however, another question remains, that is, how accurate is the contour extraction scheme? Although in the previous section it is demonstrated that the semi-automatic contour extraction scheme successfully defines the prostate contours which are a quite impressive match of the subjective contours that a human observer would infer from the original images, it is necessary to evaluate it quantitatively to ensure that its performance is reliable, better than or at least no worse than that of the manual outlining process. This will be the topic of the next chapter.
Chapter 11

Quantitative Evaluation of the Contour Extraction Scheme

The previous chapter focused on the design of the semi-automatic contour extraction scheme. Some contour delineation results generated by this scheme were presented. As we have seen, the contours identified by the scheme are visually a close match of the prostate boundaries that a human observer would infer from the original images. This suggests the possibility that it could be used as a dependable replacement or an alternative reference for the manual outlining process. It is now necessary to evaluate it quantitatively and demonstrate that its performance is reliable, better than, or at least no worse than that of the manual outlining process.

Quantitative evaluation was not addressed as part of the research in [262, 264, 265]. Focus was placed on the comparison of prostate volumes in evaluation of the computer-aided segmentation methods [264, 283, 284]. However, it was shown that prostate volume is not a suitable validation metric as different outlining by different human operators could lead to the same volume estimates, simply because volume is a cumulative function of the area enclosed by the boundaries delineated in each image [267].

In this chapter we first use distance-based and then area-based criteria to assess the segmentation results of our semi-automatic contour extraction scheme against the boundaries manually outlined by a human operator. The distance-based metrics are defined on the basis of the Hausdorff distance [357, 358] and are used to assess the reproducibility or consistency of the contour extraction scheme. The area-based metrics are used to evaluate the accuracy and sensitivity of the scheme, which are dependent on the area enclosed by the contour extracted. It will be shown that our semi-automatic contour extraction scheme surpasses the manual outline with respect to the distance-based metrics and is comparable with the manual outline in terms of the area-based criterion.
11.1 Pixel Size Calibration

To be relevant to the physical size of the organ of the prostate, the distance-based and area-based metrics will be computed in the metric units, millimetres \((mm)\) for the distance-based metrics and square millimetres \((mm^2)\) for the area-based metrics, rather than in pixels. Therefore, correspondence must be established between a pixel in the ultrasound images and the size of the physical entity it represents.

Along the margins of the ultrasound images there are calibration markers, as we have seen in Figures 2.9-2.12. The distance between two primary markers is 10\(mm\). The distances between these markers are measured manually in pixels in the images. The average distance is 51 pixels, which means each pixel in the image represents a physical entity of diameter approximately 0.2\(mm\).

11.2 Distance-based Metrics

Distance-based metrics are employed to assess the consistency or reproducibility of the contour extraction scheme against that of the manual outlining process. First, intra-observer variability is analysed for these sets of contours by measuring the discrepancy between any two contours within each set. It is hypothesised that the manual outlining process is less consistent or reproducible than our semi-automatic extraction scheme and thus is prone to larger intra-observer variability. Next inter-observer variability, that is, disagreement between a computer generated contour and manual delineation is considered and compared against disagreement between the human operators. The distance-based metrics are defined on the basis of the Hausdorff distance [357, 358].

11.2.1 The Hausdorff Distance

The Hausdorff distance is a well known metric over two closed, bounded sets and is defined as the maximum of the distances from a point in any one of the two sets to the nearest point in the other set. In the context of computer vision or pattern recognition, the two sets might represent two shapes: an “object” and a “model” against which the “object” should be matched. By translating the “model” and finding the minimum Hausdorff distance between the “object” and the “model”, the “object” will be aligned with and match the “model” as well as possible [357, 358]. Computing the Hausdorff distance between two point sets on a line as a function of translation was considered in [357] and efficient algorithms for computing the Hausdorff distance were proposed for both the case of translation and the more general case of rigid motion in [358].

We will suppose two shapes are superimposed on one another in the following text. The Hausdorff distance between them reflects the degree to which one shape resembles the other. Computing the Hausdorff distance is similar to other pattern recognition operations such as correlation and template matching, which are common practice techniques for
QUANTITATIVE EVALUATION OF THE CONTOUR EXTRACTION SCHEME

determining the degree of resemblance of two shapes [358].

The Hausdorff distance $D_H(A, B)$ between two finite point sets $A = \{a_i | 1 \leq i \leq M\}$ and $B = \{b_j | 1 \leq j \leq N\}$ is a max-min function, defined as the maximum of the two corresponding directed Hausdorff distances $D_h(A, B)$ and $D_h(B, A)$, that is,

$$D_H(A, B) = \max(D_h(A, B), D_h(B, A))$$

where the directed Hausdorff distance from $A$ to $B$ is defined as

$$D_h(A, B) = \max_i \min_j ||a_i - b_j||$$

where $|| \cdot ||$ represents some underlying norm, such as the Euclidean or $L^2$ norm. The Euclidean norm is assumed in the following text.

Equation (11.2) can be interpreted as follows. For each point $a_i \in A$, a matched “neighbour” point is identified in $B$, which is nearest to $a_i$ in terms of Euclidean distance. Note there is no explicit pairing of the points in $A$ with that in $B$. In other words, it is possible that many points in $A$ have the same closest “neighbour” point in $B$ [358]. The directed Hausdorff distance $D_h(A, B)$ hinges on the point in $A$ whose “neighbour” point in $B$ is the farthest. In other words, $D_h(A, B) = d$ means that the distance between any point in $A$ and its nearest “neighbour” in $B$ is less than or equal to $d$, and there is some point in $A$, whose distance to its closest “neighbour” point in $B$ is exactly $d$, which is the worst case in $A$, that is, the “least” matched point.

The directed Hausdorff distance from $B$ to $A$, $D_h(B, A)$ is defined similarly. The Hausdorff distance, the maximum of $D_h(A, B)$ and $D_h(B, A)$, assesses the extent to which $A$ matches $B$ provided shape $A$ lies on top of $B$. Clearly the smaller the Hausdorff distance, the more $A$ and $B$ resemble each other or become aligned with each other. The larger the Hausdorff distance, the more $A$ mismatches $B$, or in other words, the farther $A$ “deviates” from $B$. Note unlike its “standard” usage, where the primary focus is to find a best match of $A$ to $B$ between all possible relative positions by examining the Hausdorff distances, in the following text the Hausdorff distance is rather used as an indicator of the extent to which one contour “deviates” from another.

Based on the Hausdorff distance, a methodology was proposed for evaluating medical image segmentation algorithms in [359]. It was put into practice in two different ultrasound imaging applications and found to be useful in assessing the results of the segmentation algorithms designed by the researchers against the manual outlining of multiple observers.

Similar methodology was used to validate the segmentation results of cardiac boundary detection on echocardiographic sequences [360] and prostate boundary delineation in ultrasound images [267]. The computer-generated boundaries were compared against the manual outlines and it was shown that the disagreement between them was comparable to inter-observer distances of manual delineation [360] or less than inter-observer disagreement [267].
Table 11.1: Intra-observer variability: comparison of the Hausdorff distance $D_H$ of the contour extraction scheme (CES) against that of the manual outlining of two human operators (MO1 and MO2). The performance of CES is much better than manual outlining in all statistics.

<table>
<thead>
<tr>
<th>IMAGE NAME</th>
<th>MEDIAN (mm)</th>
<th>MEAN±STD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MO1</td>
<td>MO2</td>
</tr>
<tr>
<td>us1</td>
<td>1.53</td>
<td>1.91</td>
</tr>
<tr>
<td>vs01</td>
<td>1.24</td>
<td>1.54</td>
</tr>
<tr>
<td>vs04</td>
<td>1.20</td>
<td>1.54</td>
</tr>
<tr>
<td>vs12</td>
<td>1.52</td>
<td>1.53</td>
</tr>
<tr>
<td>vs18</td>
<td>1.30</td>
<td>1.45</td>
</tr>
<tr>
<td>OVERALL</td>
<td>1.34</td>
<td>1.59</td>
</tr>
</tbody>
</table>

11.2.2 Intra-Observer Variability

To compare consistency of our contour extraction scheme against that of the manual outlining, intra-observer variability is examined first. For each ultrasound image, the boundary of the prostate was manually outlined by two human operators and delineated semi-automatically by the contour extraction scheme. For each case the segmentation process was repeated ten times for all of the images. So three sets of segmentation results (ten for each set, thirty in total) are obtained for each image. For the manual outlining, a dedicated computer program was developed allowing a human operator to manually outline the prostate contour using a mouse. Each human operator consecutively outlined all of the ultrasound images (which were presented to the operator in the same order every time), took a break, and then repeatedly outlined all of the images nine more times.

Comparison of the Hausdorff Distance $D_H$

For each set of segmentation results, the Hausdorff distance $D_H(A, B)$ is computed. Here $A$ and $B$ represents any pair of segmentation results of the set. Statistics on the Hausdorff distance are collected for each set of segmentation results and results for five images are shown in Table 11.1.

The first three columns of Table 11.1 compare the sample medians of the Hausdorff distances of the three sets of segmentation results, that is, manual outlining by the first (MO1) and the second (MO2) human operator and finally the contour extraction scheme (CES). The next three columns compare the corresponding sample means and standard
deviations. The last row lists the overall sample medians, means and standard deviations.

As we can see in Table 11.1, the intra-observer variability of the two human operators (MO1 and MO2) are comparable in terms of all of the sample statistics. Take the median for instance. The overall median is 1.34 mm for MO1, which is not a far cry from that of MO2, 1.59 mm. A similar situation holds true for the overall means, which are 1.39 mm and 1.66 mm respectively for MO1 and MO2. No significant difference is expected for the standard deviation as well, ±0.36 mm for MO1 and ±0.44 mm for MO2. The means for MO1 and MO2 are within one standard deviation of each other.

However, it is a different story examining the corresponding statistics for CES in Table 11.1. It comes as no surprise that CES outperforms the manual outlining in respect of all of the statistics, both overall and that of individual images. The overall median for CES is 0.12 mm, which is less than one tenth that of MO1 and MO2, indicating a much better reproducibility or consistency.

Figure 11.1 compares the boxplots of the Hausdorff distance $D_H$ of MO1, MO2 and CES. The lower and the upper sides of the rectangles respectively correspond to the lower and upper quartiles of the data. The horizontal lines cross the rectangles at the medians. The whiskers extend to $1.5IQR$ from either end of the rectangles. The means are denoted by “×” and outliers by “+”.

$^{1}$IQR: interquartile range.
As we can see in Figure 11.1, the median and the mean of $D_H$ of CES are significantly smaller than their counterparts of MO1 and MO2. The IQR of CES is also less than that of MO1 and MO2, indicating a smaller variance. Although the differences of the sample medians and means between the two human operators are not trivial, they are not as significant as their respective differences to CES.

Moreover, the $D_H$ of MO1 and MO2 seem to have a similar distribution, which is significantly different from that of CES, as illustrated in Figure 11.2, which compares the corresponding empirical cumulative distribution functions (e.d.f.) of $D_H$. As we can see, the e.d.f. for MO1 is shifted to the left of the e.d.f. for MO2 and they do not cross each other, indicating that in terms of $D_H$ the intra-observer variability of the first human operator is smaller than that of the second human operator (this can also be inferred from the boxplots in Figure 11.1). Moreover, the two e.d.f.s are quite similar in shape, suggesting the likelihood of a similar distribution.

The e.d.f. for CES is shifted to the far left of that of MO1 and MO2 with a rather compact shape compared to the counterparts of MO1 and MO2. Its 99th percentile does not overlap with the first percentile of MO1 and MO2. This lends to support the previous hypothesis that CES has a much better reproducibility or consistency.

Next analysis-of-variance (ANOVA) was conducted to formally compare the intra-observer variability of the human operators and CES in terms of the Hausdorff distance $D_H$. The null hypothesis $H_0$ here is that the performance of MO1, MO2 and CES are identical, that is, the means of $D_H$ are the same: $\mu_{MO1} = \mu_{MO2} = \mu_{CES}$.

As we can see from Figure 11.1, the medians are less than the means and there are a number of outliers beyond the upper quartiles, which indicate a positive-skewed dis-
QUANTITATIVE EVALUATION OF THE CONTOUR EXTRACTION SCHEME

<table>
<thead>
<tr>
<th>NULL HYPOTHESIS $H_0$</th>
<th>$\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{MO1} = \mu_{MO2} = \mu_{CES}$</td>
<td>472.22</td>
<td>0</td>
</tr>
<tr>
<td>$\mu_{MO1} = \mu_{CES}$</td>
<td>336.75</td>
<td>0</td>
</tr>
<tr>
<td>$\mu_{MO2} = \mu_{CES}$</td>
<td>336.75</td>
<td>0</td>
</tr>
<tr>
<td>$\mu_{MO1} = \mu_{MO2}$</td>
<td>51.44</td>
<td>$7.364 \times 10^{-13}$</td>
</tr>
</tbody>
</table>

Table 11.2: Results of the Kruskal-Wallis test on the Hausdorff distance $D_H$ of MO1, MO2 and CES at the level of significance $\alpha = 0.001$. It is safe to conclude that the means of $D_H$ of MO1, MO2 and CES are significantly different.

null hypothesis. This suggests that it is not reasonable to assume normality for the data. A Lilliefors test [61] confirms that the underlying sample data do not have normal distributions. Therefore the Kruskal-Willis test [49, 56, 61] is used to carry out the comparison.

The Kruskal-Wallis test is a non-parametric ANOVA procedure involving overall ranking of the data. If the null hypothesis is true and the sizes of $k$ independent samples are all no less than 5, the test statistics approximately follows a $\chi^2$-distribution with $k-1$ degrees of freedom

$$\chi^2_{k-1} = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$  \hspace{1cm} (11.3)

where $n_i$ is the number of observations for each sample and $N$ is the total number of observations $N = \sum n_i$. $R_i$ is the sum of ranks for each sample. The data from the $k$ independent samples are jointly ranked from 1 (for the smallest observation) to $N$ (for the largest). The decision rule is reject $H_0$ if $\chi^2 > \chi^2_{\alpha,k-1}$, where $\alpha$ is the level of significance. For this particular case we want to compare the performance of MO1, MO2 and CES, thus $k = 3$. For each set the sample size is 45 (since there are 10 contours for each set and the number of possible pairs is 45).

The null hypothesis $H_0$ is tested at the level of significance $\alpha = 0.001$. We should reject $H_0$ if $\chi^2 > \chi^2_{0.001,2} = 13.815$ (see Appendix A5 in [49]). The test result is $\chi^2 = 472.22$ and the approximate p-value is $p = 0$ (first row of Table 11.2). Thus the test provides strong evidence that the null hypothesis should be rejected.

Next additional Kruskal-Wallis tests are conducted to make sure that the hypotheses $\mu_{MO1} = \mu_{CES}$, $\mu_{MO2} = \mu_{CES}$, and $\mu_{MO1} = \mu_{MO2}$ should all be rejected. The test results are summarised as the last three rows in Table 11.2. It is safe to conclude that the means of $D_H$ of MO1, MO2 and CES are significantly different from one and another.

A series of multiple comparisons are then conducted to provide a relative ordering of the means of MO1, MO2 and CES. The results are shown in Table 11.3. The difference of mean ranks (M.R.D.) between MO1 and MO2 is -87.94 with 95% confidence interval (C.I.) of (-131.03, -44.85), therefore it is safe to say $\mu_{MO1} < \mu_{MO2}$. After examining the remaining two rows of Table 11.3, it is safe to conclude from the above analysis that the means of $D_H$ of MO1, MO2 and CES are significantly different, and their relative ordering is $\mu_{MO2} > \mu_{MO1} > \mu_{CES}$, with CES achieving the smallest mean.
11.2. DISTANCE-BASED METRICS

<table>
<thead>
<tr>
<th>COMPARISON</th>
<th>M.R.D.</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_{MO1} - R_{MO2} )</td>
<td>-87.94</td>
<td>(-131.03, -44.85)</td>
</tr>
<tr>
<td>( R_{MO1} - R_{CES} )</td>
<td>293.53</td>
<td>(250.44, 336.62)</td>
</tr>
<tr>
<td>( R_{MO2} - R_{CES} )</td>
<td>381.47</td>
<td>(338.38, 424.56)</td>
</tr>
</tbody>
</table>

Table 11.3: Multiple comparison of difference of mean ranks (M.R.D.) of the Hausdorff distance \( D_H \) of MO1, MO2 and CES. It is safe to conclude that \( \mu_{MO2} > \mu_{MO1} > \mu_{CES} \), with CES achieving the smallest mean.

**Comparison of the Distance to the Match Point \( D_M \)**

For any pair of segmentation results \( A \) and \( B \), the distance to the match point \( D_M \) from \( A \) to \( B \) is defined as

\[
D_M(a_i, B) = \min_j ||a_i - b_j||. \tag{11.4}
\]

While the Hausdorff distance \( D_H \) measures the worst case of the extent to which one segmentation result deviates from another, the distance to the match point \( D_M \) provides an indicator of the ordinary discrepancy between the two segmentation results.

Since the distribution of \( D_M \) is unknown, we choose to compare the sample means \( \bar{D}_M \). According to the *Central Limit Theorem* [49, 56] the sampling distribution for the sample mean approximately follows a normal distribution as the sample size increases. 500 samples of \( \bar{D}_M \) each were taken for MO1, MO2 and CES. Each sample was obtained from 1000 observations of \( D_M \) randomly taken from any pair of contours.

Typical results are summarised in Table 11.4. As we can see, the mean of MO1 is 0.43mm with a 95% confidence interval (C.I.) of (0.429mm, 0.431mm), which is marginally larger than that of MO2, 0.416mm with a 95% C.I. of (0.415mm, 0.417mm). As expected, the mean of CES (0.025mm) is significantly smaller than that of the manual outlining.

Figure 11.3 compares the boxplots of the sample means \( \bar{D}_M \) (note the different scales of the \( y \)-axis of the two sub-figures). As we can see, CES has fewer outliers than MO1 and MO2. The mean of MO2 is not far from that of MO1, although it is marginally smaller. However, both of them are significantly larger than that of CES.

Next multiple comparisons are conducted and the differences of the means (M.D.) are listed in Table 11.5. As we can see, the mean of \( \bar{D}_M \) of CES is smaller than that of MO1 by 0.404mm with a 95% C.I. of (0.403mm, 0.406mm), and smaller than that of MO2 by 0.390mm with a 95% C.I. of (0.389mm, 0.392mm).

It is interesting to see that MO1 is better than MO2 in respect of the Hausdorff distance \( D_H \) (see Figure 11.1), which reflects the extreme fluctuations of the contours. However, the situation is reversed when we compare the distance to the match point \( D_M \) (see Figure 11.3), which measures the average discrepancies. Therefore it is not arbitrary to say that the performance of the two human operators are comparable. However, there is no doubt that CES is superior to both of them in terms of intra-observer variability in both of the two distance-based metrics.
QUANTITATIVE EVALUATION OF THE CONTOUR EXTRACTION SCHEME

<table>
<thead>
<tr>
<th>CONTOUR SET</th>
<th>$D_M$ (mm)</th>
<th>MEDIAN</th>
<th>MEAN</th>
<th>95% C.I. of MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>MO1</td>
<td>0.429</td>
<td>0.430</td>
<td>(0.429, 0.431)</td>
<td></td>
</tr>
<tr>
<td>MO2</td>
<td>0.415</td>
<td>0.416</td>
<td>(0.415, 0.417)</td>
<td></td>
</tr>
<tr>
<td>CES</td>
<td>0.025</td>
<td>0.025</td>
<td>(0.025, 0.025)</td>
<td></td>
</tr>
</tbody>
</table>

Table 11.4: Comparison of the statistics of the sample means $D_M$. The mean of CES is significantly smaller than that of the manual outlining, which are comparable.

<table>
<thead>
<tr>
<th>COMPARISON</th>
<th>M.D. (mm)</th>
<th>95% C.I. (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{M,MO1} - D_{M,MO2}$</td>
<td>0.014</td>
<td>(0.012, 0.016)</td>
</tr>
<tr>
<td>$D_{M,MO1} - D_{M,CES}$</td>
<td>0.404</td>
<td>(0.403, 0.406)</td>
</tr>
<tr>
<td>$D_{M,MO2} - D_{M,CES}$</td>
<td>0.390</td>
<td>(0.389, 0.392)</td>
</tr>
</tbody>
</table>

Table 11.5: Comparison of the differences of the means (M.D.) of the sample means $D_M$. The last column lists the 95% C.I. of M.D..

Figure 11.3: Comparison of the boxplots of the sample means $D_M$ (note the different scales of the y-axis of the two sub-figures). The mean of CES is significantly smaller than that of MO1 and MO2, which are not a far cry from each other. CES has fewer outliers than MO1 and MO2.
11.2.3 Inter-Observer Variability

In the previous subsection, CES was compared to manual outlining with respect to the intra-observer variability. Focus there was placed on the disagreement between the contours produced by the same human operator (or CES). As we can see, the prostate boundaries from two consecutive segmentations by the same human operator may be quite different, as reflected by the large value of the Hausdorff distance based metrics, whereas the CES is more consistent in producing the “settled” contour of the prostate.

Next the inter-observer variability (disagreement between the two human operators and that between CES and a human operator) will be examined, that is, how the contours of one set of segmentation results differ from those of another. The disagreement between contours of the two human operators will be compared against the discrepancy between contours of CES and one of the human operators with respect to the Hausdorff distance $D_H$ and the distance to the match point $D_M$.

Comparison of the Hausdorff Distance $D_H$

Typical results of the Hausdorff distance $D_H$ are summarised in Table 11.6. The first three columns respectively list the medians of $D_H$ between the contours of the two human operators (denoted by MO1#MO2), between CES and the first human operator (MO1#CES), and between CES and the second human operator (MO2#CES). The last three columns are the corresponding means and standard deviations. The overall statistics are shown in the last row.

As we can see in Table 11.6, the average discrepancy between CES and the first human operator (1.80 mm) is less than that between the two human operators (2.59 mm), which is comparable to that between CES and the second human operator (2.54 mm). An additional Kruskal-Wallis test reveals that $\mu_{MO1#MO2} = \mu_{MO2#CES}$ cannot be rejected since $\chi^2 = 0.61$ and $p$-value $p = 0.4333$ and a multiple comparison shows that the 95% C.I. of their difference of mean ranks contains zero.

Figure 11.4 compares the boxplots of the Hausdorff distance $D_H$ between contours of the two human operators (MO1#MO2), between the first human operator and the CES (MO1#CES), and between the second human operator and the CES (MO2#CES). As we can see, the discrepancies between CES and either one of the two human operators are no more than that between the two human operators. The IQRs of MO1#CES and of MO2#CES are all smaller than that of MO1#MO2, indicating a smaller variance, which is attributable to the smaller intra-observer variability of CES.

Comparison of the Distance to the Match Point $D_M$

As before we choose to compare the sample means $\overline{D_M}$. 500 samples of $\overline{D_M}$ each were taken for MO1#MO2, MO1#CES and MO2#CES. Each sample was obtained from 1000 observations of $D_M$ randomly taken from any pair of contours.
Table 11.6: Inter-observer variability: comparison of the Hausdorff distance $D_H$ between the contours of the two human operators (MO1#MO2), between the first human operator and CES (MO1#CES), and between the second human operator and CES (MO2#CES). The average discrepancy of MO1#CES is less than that of MO1#MO2, which is comparable to that of MO2#CES.

![Figure 11.4: Inter-observer variability: comparison of the boxplots of the Hausdorff distance $D_H$ between the contours of the two human operators (MO1#MO2), between the first human operator and CES (MO1#CES), and between the second human operator and CES (MO2#CES). The rectangles represent the central halves of the data. The horizontal lines cross the rectangles at the medians. The whiskers extend to $1.5IQR$ from either end of the rectangles. The means are denoted by “×” and outliers by “+”.

245
11.3. AREA-BASED METRICS

<table>
<thead>
<tr>
<th>INTER-SET COMPARISON</th>
<th>$D_M (\text{mm})$</th>
<th>MEDIAN</th>
<th>MEAN</th>
<th>95% C.I. of MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>MO1 vs MO2</td>
<td>0.578</td>
<td>0.578</td>
<td></td>
<td>(0.577, 0.579)</td>
</tr>
<tr>
<td>MO1 vs CES</td>
<td>0.573</td>
<td>0.574</td>
<td></td>
<td>(0.572, 0.575)</td>
</tr>
<tr>
<td>MO2 vs CES</td>
<td>0.673</td>
<td>0.673</td>
<td></td>
<td>(0.671, 0.674)</td>
</tr>
</tbody>
</table>

Table 11.7: Comparison of medians and means of the sample means $D_M$. The mean of sample means $D_M$ of MO1 vs MO2 is comparable to that of MO1 vs CES. The mean of $D_M$ of MO2 vs CES is larger than that of the other two pairs by about 0.1 mm.

Table 11.7 compares the medians and means of the sample means $D_M$. As we can see, the mean of $D_M$ of MO1 vs MO2 is estimated to be 0.578 mm with a 95% C.I. of (0.577 mm, 0.579 mm), which is comparable to that of MO1 vs CES, 0.574 mm with a 95% C.I. of (0.572 mm, 0.575 mm). The mean of $D_M$ of MO2 vs CES is 0.673 mm, larger than that of the other two pairs by about 0.1 mm.

The corresponding boxplots of the sample means $D_M$ are shown in Figure 11.5. As expected, the boxplots of MO1 vs MO2 and MO1 vs CES are very similar, although the median and mean of MO1 vs CES are marginally smaller. As we can see, the central half of MO2 vs CES is about 0.1 mm above that of MO1 vs CES, indicating that on the average the contours of CES agree more with that of MO1 than with that of MO2.

The discrepancy between CES and the human operators (especially MO2 vs CES) could rather be explained by their different tendencies in identifying the prostate boundary. For instance, in some areas, especially where the boundaries are vague, what CES locates may be very different from what a human operator interprets. This leads to the discussion of the accuracy and sensitivity of CES in the next section.

11.3 Area-based Metrics

In the previous section CES was assessed in terms of intra-observer and inter-observer variability based on the Hausdorff distance and the distance to the match point. Distance-based metrics provide local measures of the difference between contours and therefore are sensitive to local shape changes. In contrast, area-based metrics are global measures and insensitive to shape. To assess the accuracy and sensitivity of the CES, we need to compare a contour extracted by CES with the true contour of the prostate to find out the common and different areas enclosed by them. Unfortunately no ground truth contour of the prostate is available except the two sets of manual segmentation results.

However, as we have seen in the previous subsection, the intra-observer variability of manual outlining is not trivial. Therefore it is inappropriate to randomly select one contour of manual outlining and compare it with a contour of CES. Instead, it is a sensible approach to compare the CES contours against the “average” contour of the manual outline, which could be regarded as the pseudo ground truth contour of the prostate.
11.3.1 The “Average” Contour

The “average” contour of each set of segmentation results is obtained as follows.

- First the approximate common “centroid” of the area enclosed by the contour is found on the basis of the coordinates of the component points of the contours.

- Next each contour is interpolated according to the distance between each point and the “centroid” and the radian angle of the corresponding radial line emanating from the “centroid” and passing through the point.

- Finally, the coordinates of the points at the same angles are averaged over all of the interpolated contours to give the “average” contour.

As an example, the “average” contours (shown as solid lines) of the two sets of manual outlining (dotted lines) of image “us1” are shown in Figure 11.6. For comparison, the “average” contour of CES is also shown. As we can see, the discrepancy among the segmentation results of manual outlining MO1 (Figure 11.6(a)) is “uniform” and “universal” along the whole contour, as characterised by the “well-distributed” contour points around
Figure 11.6: The “average” contours (solid lines) of the three sets of segmentation results (dotted lines) of image “us1”. Note the “universal” discrepancy among the segmentation results of manual outlining (MO1 and MO2), whereas for CES it is mostly confined to a handful of small segments where the boundaries of the prostate are rather obscure.
the “average” contour, no matter whether the underlying boundaries are prominent or obscure. There are very few segments where all of the segmentation results agree with each other. As expected, the same situation holds true for MO2 (Figure 11.6(b)), which confirms the large intra-observer variability of the manual outlining.

Note the considerable deviation from the “average” contour at the lower right corner of MO1, which puts a big question mark on the credibility of the result of a single manual segmentation. The “average” contour, which could be regarded as the pseudo ground-truth contour, is more credible than the result of a single manual segmentation. Unfortunately, repeated manual delineation of the same image would not be possible in practice.

In contrast, the disagreement between the CES contours (Figure 11.6(c)) are trivial and mostly confined to a handful of small segments where the boundaries of the prostate are rather vague. In other words, although there are some “inconsistent” parts (this “inconsistency” is negligible compared to that of manual delineation), most parts of the contour are “consistent”.

Once the “average” contours of the manual outlining are available, we can validate CES in terms of accuracy and sensitivity, which are defined on the basis of the common and different areas enclosed by the “average” contour of the manual outlining and the contours extracted by CES.

11.3.2 Definition of the Accuracy \( M_a \) and Sensitivity \( M_s \)

Figure 11.7 illustrates a typical scenario with the “average” contour of the manual outlining (shown as dotted line) and the corresponding contour extracted by CES (shown as solid line) superimposed on each other. The different regions enclosed by the two curves are defined respectively as:

- **True positive (TP):** the common (“intersection”) region enclosed both by the “average” contour of manual outline and by the contour extracted by CES.

- **False positive (FP):** the region covered by the CES contour only, not enclosed by the “average” contour of manual outline.

- **False negative (FN):** the region enclosed by the “average” contour of manual outline only, not covered by the CES contour.

- **True negative (TN):** the region neither covered by the “average” contour of manual outline nor by the CES contour.

If the area of the first three regions above are respectively denoted by \( A_{TP} \), \( A_{FP} \), and \( A_{FN} \), then the area enclosed by the “average” manual outline is

\[
A_{MOA} = A_{TP} + A_{FN}
\]
11.3. AREA-BASED METRICS

Figure 11.7: Definition of the different regions enclosed by the “average” contour (dotted line) of the manual outlining and the contour extracted by the CES (solid line): true positive (TP), false positive (FP) and false negative (FN). The region neither covered by the “average” manual contour nor by the CES contour is denoted by true negative (TN). The accuracy and sensitivity of the CES are defined on the basis of the area of these regions.

and the area enclosed by either the “average” manual outline or the contour of CES (that is, the “union” of them) is

\[ A_U = A_{TP} + A_{FN} + A_{FP}. \]  

(11.6)

Based on these areas the sensitivity [270] of the CES is defined as

\[ M_s = \frac{A_{TP}}{A_{MOA}}. \]  

(11.7)

and the accuracy as

\[ M_a = 1 - \frac{A_{FP} + A_{FN}}{A_{MOA}}. \]  

(11.8)

After considering equation (11.6) the above becomes

\[ M_a = 1 - \frac{A_U - A_{TP}}{A_{MOA}}. \]  

(11.9)

Therefore, in practice only \( A_{MOA} \), \( A_U \) and \( A_{TP} \) need to be determined and they are computed using Green’s theorem\(^2\).

The area \( A_{MOA} \) can be easily determined since the coordinates of the contour points are already available (refer to previous subsection on obtaining the “average” contour).

\(^2\) According to Green’s theorem, the area enclosed by a closed curve \( C \) in the \( x, y \)-plane is given by the contour integration carried out around the closed curve

\[ A = \frac{1}{2} \oint_C (xdy - ydx). \]  

(11.10)
QUANTITATIVE EVALUATION OF THE CONTOUR EXTRACTION SCHEME

<table>
<thead>
<tr>
<th>IMAGE NAME</th>
<th>$M_s$ (%): MOA1</th>
<th>$M_s$ (%): MOA2</th>
<th>$M_s$ (%): MOA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MEDIAN</td>
<td>MEAN</td>
<td>MEDIAN</td>
</tr>
<tr>
<td>us1</td>
<td>95.31</td>
<td>95.32</td>
<td>94.54</td>
</tr>
<tr>
<td>vs01</td>
<td>92.07</td>
<td>92.08</td>
<td>92.43</td>
</tr>
<tr>
<td>vs04</td>
<td>95.27</td>
<td>95.26</td>
<td>94.41</td>
</tr>
<tr>
<td>vs12</td>
<td>90.23</td>
<td>90.34</td>
<td>83.77</td>
</tr>
<tr>
<td>vs18</td>
<td>96.35</td>
<td>96.36</td>
<td>94.82</td>
</tr>
<tr>
<td>OVERALL</td>
<td>95.24</td>
<td>93.87</td>
<td>94.40</td>
</tr>
</tbody>
</table>

Table 11.8: Assessment of CES by area-based metrics: sensitivity $M_s$. Most of the time $M_s$ is over 90%. The overall median and mean of $M_s$ are respectively 95.07% and 93.33% with MOA as the pseudo ground truth contour.

However, this is not the case for $A_U$ and $A_{TP}$. As we can see in Figure 11.7, the contour bordering $A_U$ and $A_{TP}$ comprises both “MOA parts” and “CES parts”. For computation of $A_U$ and $A_{TP}$, the corresponding component contour points need to be identified first. This is done as follows.

- First, the common “centroid” of the “average” contour of the manual outlining and the contour of CES is obtained by averaging the coordinates of the two contours.

- Next the two contours are interpolated on the basis of the distance between each contour point and the “centroid” and the corresponding radian angle of that point.

- Finally, at each radian angle the two contour points are compared according to their distance to the “centroid” and the innermost one is deemed to belong to the contour enclosing $A_{TP}$ and the outermost one as enclosing $A_U$ (refer to Figure 11.7).

After identifying the contour points bordering $A_{TP}$ and $A_U$, determining these two areas is straightforward according to Green’s theorem.

11.3.3 Assessment of the Accuracy $M_a$ and Sensitivity $M_s$ of CES

Each CES contour is compared with the “average” manual outline of the first human operator (denoted by MOA1), with that of the second human operator (MOA2), and with the overall “average” manual outline of the two human operators (MOA). Typical results of assessment of CES in terms of sensitivity $M_s$ are shown in Table 11.8.

The first two columns of Table 11.8 respectively show the medians and means of the sensitivity $M_s$ of CES, with MOA1 as the pseudo ground truth contour. The last two columns whose discrete form is

$$A = \frac{1}{2} \sum_{i=1}^{N} [x_i(y_{i+1} - y_i) - y_i(x_{i+1} - x_i)]$$  \hspace{1cm} (11.11)

which reduces to

$$A = \frac{1}{2} \sum_{i=1}^{N} (x_i y_{i+1} - x_{i+1} y_i)$$  \hspace{1cm} (11.12)

where $N$ is the number of points of the discrete contour and $(x_i, y_i)$ is the coordinate of the $i$-th point of the contour [117].
11.3. AREA-BASED METRICS

Table 11.9: Validation of CES by area-based metrics: accuracy $M_a$. The overall median and mean of $M_a$ are respectively 94.05% and 92.67% with MOA as the pseudo ground truth contour. Most of the time $M_a$ is over 90%.

<table>
<thead>
<tr>
<th>IMAGE NAME</th>
<th>$M_a$(%): MOA1</th>
<th>$M_a$(%): MOA2</th>
<th>$M_a$(%): MOA</th>
</tr>
</thead>
<tbody>
<tr>
<td>us1</td>
<td>95.06</td>
<td>95.05</td>
<td>94.61</td>
</tr>
<tr>
<td>vs01</td>
<td>91.77</td>
<td>91.78</td>
<td>92.25</td>
</tr>
<tr>
<td>vs04</td>
<td>94.27</td>
<td>94.25</td>
<td>94.50</td>
</tr>
<tr>
<td>vs12</td>
<td>88.53</td>
<td>88.64</td>
<td>87.79</td>
</tr>
<tr>
<td>vs18</td>
<td>94.04</td>
<td>94.03</td>
<td>94.05</td>
</tr>
<tr>
<td>OVERALL</td>
<td>94.04</td>
<td>92.75</td>
<td>94.05</td>
</tr>
</tbody>
</table>

columns are the results with MOA as the pseudo ground truth contour. As we can see, most of the time the sensitivity of CES is over 90%. The overall median and mean of the sensitivity are respectively 95.07% and 93.33% with MOA as the pseudo ground truth contour. The worst case is that of image “vs12” with MOA2 as the pseudo ground truth contour, 83.87% in average, which, however, is not far below 90%.

It is a similar story for assessment of the accuracy $M_a$, as shown in Table 11.9. Most of the time the accuracy is over 90%. The overall median and mean of the accuracy are respectively 94.05% and 92.67% with MOA as the pseudo ground truth contour. By comparing Table 11.8 and Table 11.9, we can see that $M_a$ is always marginally lower than the corresponding $M_s$. This is actually due to the definition of the accuracy and sensitivity. Figure 11.8 illustrates a typical scenario when assessing the sensitivity and accuracy of CES, in which the “average” manual outline MOA1 (dashed line) is superimposed on a contour extracted by CES (solid line) for image “us1”.

As we can see, most of the CES contour lies inside MOA1 (but very close to it), which accounts for the high sensitivity of the CES contour. Although in this case the “error” of the CES contour can be mostly attributed to the false negative (FN) areas rather than the false positive (FP) areas, it is the FP areas that make the accuracy of the CES contour smaller than the corresponding sensitivity. As we can see from equation (11.7-11.9) and Figure 11.8, if there were no FP areas (that is, the CES contour lay entirely inside MOA1), the sensitivity and the accuracy would be equal (as an example, $M_s$ and $M_a$ for image “vs12” are equal with MOA2 or MOA as the pseudo ground truth contour).

In other words, the sensitivity only depends on the TP areas that are correctly claimed as the prostate region by CES. For the accuracy, we are also concerned about the FP areas which are falsely claimed by CES and penalise them. The larger the FP areas, the smaller the accuracy. Ultimately, we are more concerned about the accuracy rather than the sensitivity.

Since each human operator is subject to his own bias when manually delineating the prostate contour, it is fair to say that the overall “average” manual outline MOA should be trusted (rather than the individual one MOA1 or MOA2) when evaluating the accuracy
QUANTITATIVE EVALUATION OF THE CONTOUR EXTRACTION SCHEME

Figure 11.8: A typical scenario when assessing the sensitivity and accuracy of CES for image “us1”. The “average” contour of manual outline is shown as dashed line and the CES contour as solid line. Note it is the FP area that makes the accuracy of the CES smaller than the corresponding sensitivity.

of CES.

Figure 11.9 and Figure 11.10 respectively compare the boxplots of the sensitivity $M_s$ and accuracy $M_a$ of CES with that of the two human operators. The overall “average” manual outline MOA is used as the pseudo ground truth contour. The outliers are due to image “vs12” for which the different tendencies of CES and the human operators (especially the second operator) in identifying the prostate boundary are not trivial (with CES on the conservative side, as the CES contours lie entirely inside the manual outlined contours). Nevertheless, as we can see in Figure 11.10, the accuracy of CES is comparable to that of the two human operators. Therefore we can conclude that it is highly reliable in locating the prostate boundary.

11.4 Conclusion

Before our semi-automatic contour extraction scheme can be used as a dependable replacement or an alternative reference for the manual outlining process, we have to be assured that its performance is more reliable than that of the manual outlining. Accordingly, our focus in this chapter is placed on quantitative assessment of the semi-automatic contour extraction scheme against the manual outlining.

As we have seen, CES is more consistent than a human operator in terms of “intra-observer” variability and comparable with respect to inter-observer variability. Note all of the manual outlining was conducted with great care and patience, and most importantly, without time constraint or any other forms of pressure. In practice, however, this would not be possible with so many images for the radiologist to interpret everyday, since repeated
manual outlining is tedious, challenging and time-consuming. Therefore it is not arbitrary to say that greater intra-observer and inter-observer variability can be expected in day to day clinical practice. CES is superior to the manual outlining in this sense.

Obviously, reproducibility of CES is much better than that of the manual outlining in terms of intra-observer variability, however, there is still notable “inconsistency” between two CES generated boundaries. Ideally CES should generate exactly the same contour every time without any variability. Why does this inconsistency arise? It can be explained as follows.

Although the parameters of the snake are always kept the same, an unavoidable fact is that the result of the seed point identification and snake initialisation process are different every time, simply because it is unlikely for the user to identify the anchor point exactly at the same position every time. Thus different anchor points result in slightly different seed points.

As we know from the previous chapters, the “strength” of external forces are stronger around prominent boundaries but weaker around relatively feature free areas. Therefore, the parts of a snake which are close to prominent boundaries tend to move quickly towards them because of the underlying strong external forces, no matter how “far” they
threshold initially are from the boundaries (of course within a reasonable range). Whereas the parts initialised near a “flat” area are more likely to stay where they initially are due to lack of external forces. In other words, it does not matter where the snake is initialised for the areas with prominent boundaries, the snake will eventually settle down on to the boundaries. Therefore, there is no notable inconsistency for these areas. However, initialisation does matter for the relatively feature free areas, which accounts for the inconsistency of the snake.

Despite the negligible inconsistency of CES (compared to the large intra-observer variability of manual delineation), the validation results reveal that the contours generated by the computer are reliable and far more consistent than that of manual delineation, which is more often than not subject to a radiologist’s expertise on one hand and bias on the other hand. Therefore, it is not arbitrary to say that our semi-automatic contour extraction scheme is reliable and could be used as a dependable replacement or a highly accurate alternative reference for the manual outlining process.
11.4. CONCLUSION
Chapter 12

Conclusion

Of relevance to various issues about computer assisted analysis of medical images in general, this thesis focuses on multiresolution analysis of ultrasound images of the prostate. Concretely, from Chapter 3 to Chapter 11 in turn we deal with the following tasks:

- ultrasound image modelling,
- ultrasound speckle reduction, and
- prostate contour extraction,

which respectively fit in the spectrum of computer aided analysis of medical images addressing medical imaging modality, preprocessing, and segmentation (see Figure 1.1). Emphasis has been placed on taking a multiresolution perspective throughout the thesis, either in suppressing ultrasound speckle or in extracting the contour of the prostate.

12.1 Contributions of Thesis

Specifically, the main contributions of this thesis are summarised as follows.

- Computer simulation of B-mode ultrasound imaging, which not only provides us with an insight into the nature of speckle, but also serves a viable test-bed for our speckle reduction scheme or the like. The computer simulation of B-scan ultrasound imaging and the subsequent analysis of the statistical properties of the simulated images inspired us to empirically propose the generalised Fisher-Tippett distribution which best fits the ultrasound images of the prostate.

- A new speckle noise reduction scheme in the wavelet transform domain, which is on the basis of modelling the statistical properties of the wavelet coefficients, and takes advantage of their inter-scale correlation. The squared modulus of the component wavelet coefficients were modelled as a two-state Gamma mixture. Inter-scale correlation was exploited by taking the harmonic mean of the posterior probability functions, which are derived from the Gamma mixture. This noise reduction scheme
was applied to both simulated and real ultrasound images, and its performance is quite satisfactory in that the important features of the original noise corrupted image are preserved while most of the speckle noise is removed successfully. It was also evaluated both qualitatively and quantitatively by comparing it with median, Wiener, and Lee filters, and the results revealed that it definitely surpasses all these filters in comparison.

- A novel contour extraction scheme (CES), which fuses the dyadic wavelet transform (DWT) and snakes. The problem of extracting the contour of the prostate was placed in a multi-scale framework provided by the DWT. Specifically, the external potential functions of the snake are designated as the modulus of the wavelet coefficients at different scales provided by the DWT, and thus are “switchable”. Such a multi-scale snake, which deforms and migrates from coarse to fine scales, eventually extracts the contour of the prostate. The CES was evaluated quantitatively using both distance-based and area-based criteria, and it was found that the CES surpasses the manual outline with respect to the distance-based metrics and is comparable with the manual outline in terms of the area-based criterion. This indicates that the CES is reliable and could be used as a dependable replacement or a highly accurate alternative reference for the subjective and hardly reproducible manual outlining process.

Other nontrivial contributions of the thesis are:

- a thorough up-to-date literature review of the speckle distribution models, speckle reduction methods, and prostate contour delineation;
- a concise introduction to the concepts of multiresolution analysis in general, and Mallat and Zhong’s dyadic wavelet transform in particular; and
- an accessible elaboration of deformable models and their implications for the computer aided medical image analysis. The mathematical foundation, namely Euler’s equations for snakes and dynamic deformable models are also provided for completeness.

12.2 Implications and Suggestions for Future Work

There is much room left for further improvement or extension. Suggestions for future work are listed below.

- For B-mode ultrasound imaging, our computer simulation is a simplified version of a real world ultrasonic imaging system. To make it more realistic, some assumptions can be changed or more imaging factors can be introduced into the simulation process. For instance, a non-Gaussian lateral beam profile with sidelobes, which is appropriate to a transducer with a rectangular aperture, can be simulated approximately by a sinc function.
CONCLUSION

- As to statistical properties of speckled images, as we have seen in Chapter 4, the Rayleigh distribution is adequate for modelling speckle of a homogeneous phantom. Although various distributions have been tried to model phantoms of complex structures, it is evident that a single distribution model is inappropriate for describing all possible scenarios of speckled ultrasound images. It is therefore worthwhile to explore a mixture of distribution models, which can be of the same type or different ones.

- For speckle reduction, it is a common practice to formulate it as an estimation problem where the aim is to recover the original image from its noisy observation by following some criteria such as minimum mean square error (MMSE) or maximum a posteriori (MAP). It is not surprising to see that the same criteria are adopted in designing filtering methods in the wavelet domain. Since the wavelet transform partially decorrelates the noise and there are more than one type of wavelets, many possibilities are there. Another possibility is to exploit inter-scale and intra-scale dependency of the wavelet coefficients by resorting to Markov random field (MRF) models or hidden Markov trees.

- As to prostate contour extraction, dynamic deformable models can be incorporated into the process to track the prostate contour in image sequences. A dynamic snake also makes it possible to introduce user’s intervention in real time, which is sometimes beneficial.

The above suggestions are specific to the thesis proper. There is much to explore beyond the scope of the thesis as well. Referring back to Figure 1.1, we list some possibilities as follows.

- After the prostate contour is available, the region of interest can be easily defined. It is then possible to investigate problems such as tissue characterisation, cancer (hypoechoic) or calcification (hyperechoic) detection, within the region of the prostate.

- As to quantitative analysis, we actually have given an example in evaluation of the CES in section 11.3, where using Green’s theorem we calculate cross-sectional areas of the prostate from the contours extracted. Another possibility is to derive the volume of the prostate or examine its symmetry on the basis of the contours of the prostate.

- After contours of the prostate are extracted from sequence of images (which are obtained by moving the probe from the cephalad to the caudad extent of the prostate), it is possible to derive a 3D visualisation of the prostate from them, which has a strong implication for surgical or radiation therapy planning such as brachytherapy.

It is also hoped that the multiscale-snake-based contour extraction scheme in Chapter 10, particularly the semi-automatic centre-based approach for snake initialisation find appli-
cation in other medical images, where the objects of interest approximately have the shape of a sphere, spheroid, or ellipsoid.
Appendix A

Euler’s Equation for Snakes

For the sake of completeness, Euler’s equation for snakes is presented briefly in this appendix. Determining the minimum of the energy functional of snakes (9.8) falls into a wide class of problems that is the subject of the Calculus of Variations [310, 311], where the objective is to determine extrema or stationary values for functionals.

A.1 Functionals

A functional is different from an ordinary function in that it is a quantity or function defined by the entire course of one or more functions rather than by a few variables. In other words, its arguments are functions rather than a number of variables and its domain is a space of functions rather than a region in coordinate space. For example, the length $L$ of a curve $y = y(x)$ joining two points $(x_0, y_0)$ and $(x_1, y_1)$ in a plane is given by the integral

$$L(y) = \int_{x_0}^{x_1} \sqrt{1 + y'^2} \, dx \quad (x_0 < x_1)$$

(A.1)

This is a functional which depends on the argument function $y(x)$.

A.2 Fundamental Lemma of the Calculus of Variations

Lemma 1 If a function $\phi(x)$ is continuous on $[x_0, x_1]$ and satisfies

$$\int_{x_0}^{x_1} \eta(x) \phi(x) \, dx = 0$$

(A.2)

for all functions $\eta(x)$ which vanish on the boundary ($\eta(x_0) = 0, \eta(x_1) = 0$) and are continuous together with their first two derivatives, then it follows that $\phi(x) \equiv 0$ on $[x_0, x_1]$.

Proof: This lemma is easily proved by contradiction. Suppose on the contrary that $\phi(x) > 0$ at $x = \xi$. Since $\phi(x)$ is continuous, there must exist a neighbourhood $N$
A.3. **THE EULER EQUATIONS**

containing $\xi$, given by $\xi_0 < \xi < \xi_1$, in which $\phi(x) > 0$. Note that (A.2) holds for any function $\eta(x)$, without loss of generality and for the sake of simplicity, we can define $\eta(x)$ by

$$
\eta(x) = \begin{cases} 
0 & \text{for } x_0 \leq x \leq \xi_0 \\
(x - \xi_0)^4(x - \xi_1)^4 & \text{for } \xi_0 \leq x \leq \xi_1 \\
0 & \text{for } \xi_1 \leq x \leq x_1
\end{cases}
$$

Clearly $\eta(x)$ satisfies the required condition, that is, it is continuously differentiable and $\eta(x_0) = 0, \eta(x_1) = 0$. Since $\phi(x) > 0$ and $\eta(x) > 0$ in $(\xi_0, \xi_1)$ it is clear that

$$
\int_{x_0}^{x_1} \eta(x)\phi(x)dx = \int_{\xi_0}^{\xi_1} \phi(x)(x - \xi_0)^4(x - \xi_1)^4dx > 0
$$

Therefore hypothesis (A.2) is violated and we have a contradiction. A similar contradiction is reached if $\phi(x) < 0$, which proves the lemma. It is called the fundamental lemma of the Calculus of Variations or the lemma of du Bois-Reymond [310,311].

A.3 **The Euler Equations**

Let us first consider the simplest problem of the calculus of variations, that is, finding the minimum of an integral of the following form

$$
I = \int_{x_0}^{x_1} F(x, y, y')dx \tag{A.3}
$$

where $x_0, x_1, y_0 = y(x_0), y_1 = y(x_1)$ are given. The function $F$ is assumed to be twice continuously differentiable with respect to its three arguments and $y''$ is also assumed continuous.

Suppose $y = f(x)$ is the desired function which makes the integral $I$ stationary, that is, yielding the minimum. We construct a function $y(x) = f(x) + \epsilon\eta(x) = y + \delta y$, where $\eta(x)$ is any function which is defined in the interval $[x_0, x_1]$ with a continuous second derivative and which vanishes at the end points ($\eta(x_0) = 0, \eta(x_1) = 0$). $\epsilon$ is a parameter independent of $x$. ($\delta y = \epsilon\eta(x)$ is known as the variation of the function $y$.) Substitution of this function into (A.3) gives

$$
I = \int_{x_0}^{x_1} F(x, f + \epsilon\eta, f' + \epsilon\eta')dx \tag{A.4}
$$

The above integral $I$ now may be regarded as a function of $\epsilon$, that is $I = I(\epsilon)$. It is permissible that the function $I(\epsilon)$ be differentiated with respect to $\epsilon$ under the integral sign, so it follows that

$$
I'(\epsilon) = \int_{x_0}^{x_1} \left[ \frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y'} \eta' \right] dx \tag{A.5}
$$

Since $y = f(x)$ is the extremal function yielding the minimum of $I$, it immediately follows
that $I(\epsilon)$ must take a minimum at $\epsilon = 0$. Therefore

$$I'(0) = \int_{x_0}^{x_1} \left[ \frac{\partial F}{\partial y} \eta + \frac{\partial F}{\partial y'} \eta' \right] dx = 0 \quad (A.6)$$

Evaluating the second term of the integral after integrating by parts and noting that $\eta(x_0) = 0, \eta(x_1) = 0$ we obtain

$$\int_{x_0}^{x_1} \frac{\partial F}{\partial y} \eta' dx = \frac{\partial F}{\partial y} \eta \bigg|_{x_0}^{x_1} - \int_{x_0}^{x_1} \eta \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) dx$$

$$= - \int_{x_0}^{x_1} \eta \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) dx \quad (A.7)$$

Therefore (A.6) becomes

$$I'(0) = \int_{x_0}^{x_1} \eta \left[ \frac{\partial F}{\partial y} - \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) \right] dx = 0 \quad (A.8)$$

which is valid for an arbitrary function $\eta$. From the fundamental lemma of the Calculus of Variations above it immediately follows that

$$\frac{\partial F}{\partial y} - \frac{d}{dx} \left( \frac{\partial F}{\partial y'} \right) = 0 \quad (A.9)$$

or for the sake of simplicity, after a change of symbol

$$[F]_y = F_y - \frac{d}{dx} F_{y'} = 0 \quad (A.10)$$

where the subscripts indicate partial derivatives with respect to the variables concerned. The differential expression $[F]_y$ is called the variational derivative of $F$ with respect to $y$. Its role is analogous to that of the gradient in ordinary minimum problems. (A.10) is Euler’s equation [310, 311]. It is an ordinary differential equation of the second order and its validity is a necessary condition for the existence of an extremum. Therefore the variational problem of determining the minimum of the integral in (A.3) is reduced to finding a function which satisfies the Euler equation subject to the boundary conditions $y_0 = y(x_0), y_1 = y(x_1)$.

Consider the problem of finding the shortest curve between two points in a plane. From (A.1) apparently

$$F(x, y, y') = \sqrt{1 + (y')^2}$$

and the corresponding partial derivatives are

$$F_y = 0$$

$$F_{y'} = \frac{y'}{\sqrt{1 + (y')^2}}$$

263
Substitution of the above two equations into (A.10) gives the following Euler’s equation

\[ \frac{d}{dx} \left( \frac{y'}{\sqrt{1+(y')^2}} \right) = 0 \]

Integrating this differential equation yields

\[ \frac{y'}{\sqrt{1+(y')^2}} = C \]

where \( C \) represents a constant. Solving for \( y' \) gives \( y' = \text{constant} \), which means that a straight line is the shortest curve passing through two points in a plane.

### A.4 Euler’s Equations for Variational Problems with Higher Derivatives

Euler’s equations are obtained in an analogous manner for variational problems with higher derivatives, that is, finding the stationary values of the integral

\[ I = \int_{x_0}^{x_1} F(x, y, y', y'', \cdots, y^{(n)}) dx. \tag{A.11} \]

Euler’s equations are given as the following without proof (the results are easily proved in the same way as Euler’s equation above)

\[ [F]_y = F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} - \cdots + (-1)^n \frac{d^n}{dx^n} F_{y^{(n)}} = 0 \tag{A.12} \]

where \( F \) is a function of the arguments \( x, y, y', y'', \cdots, y^{(n)} \) and all those functions have continuous derivatives up to the \( 2n \)-th order for which the values are given at the boundary up to the \( (n-1) \)-th order. For \( n=2 \), which is the case for the snakes problem, we have

\[ [F]_y = F_y - \frac{d}{dx} F_{y'} + \frac{d^2}{dx^2} F_{y''} = 0 \tag{A.13} \]

### A.5 Euler’s Equation for Snakes

We rewrite the energy functional of the snakes as follows

\[ E_{\text{snake}}(v(s)) = \int_0^1 \frac{1}{2} \left( \alpha(s)|v'(s)|^2 + \beta(s)|v''(s)|^2 \right) + P(v(s)) ds \tag{A.14} \]

so

\[ F(s, v, v', v'') = \frac{1}{2} \left( \alpha|v'|^2 + \beta|v''|^2 \right) + P(v) \tag{A.15} \]
We have

\[ F_v = \nabla P(v) \]
\[ F_{v'} = \alpha v' \]
\[ F_{v''} = \beta v'' \]  \hspace{1cm} (A.16)

so the corresponding Euler’s equation is

\[ \frac{d}{ds} (\alpha v') - \frac{d^2}{ds^2} (\beta v'') - \nabla P(v) = 0 \]  \hspace{1cm} (A.17)

If both \( \alpha \) and \( \beta \) are constants, this reduces to

\[ \alpha v'' - \beta v''' - \nabla P(v) = 0 \]  \hspace{1cm} (A.18)

which cannot be explicitly integrated and thus a direct solution cannot be obtained.
Appendix B

Lagrange’s Equations of Motion for Dynamic Deformable Models

Hamilton’s principle is one of the most important concepts of mathematical physics, which leads directly to Lagrange’s equations, which in turn may be used to establish Newton’s second law [310, 311]. Lagrange’s equations can be applied to mechanical and non-mechanical systems such as electric networks. The dynamic deformable models described in section 9.5 are based on the Lagrange’s equation of motion. For the sake of completeness we briefly introduce the background of Hamilton’s principle and Lagrange’s equations of motion in this appendix.

B.1 Euler’s Equations for Variational Problems with Several Unknown Functions

We first need to establish Euler’s equations for variational problems with more than one unknown functions, that is, determining several functions \( y(x), z(x), \cdots \) of \( x \) which make the following integral
\[
I = \int_{x_0}^{x_1} F(x, y, z, \cdots, y', z', \cdots) \, dx
\]
yield an extremum. The corresponding Euler’s differential equations are given as the following without proof (as they are easily proved in a manner similar to the proof in Appendix A)
\[
[F]_y = F_y - \frac{d}{dx} F'y = 0
\]
\[
[F]_z = F_z - \frac{d}{dx} F'z = 0
\]
\[
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdott
B.2 Hamilton’s Principle and Lagrange’s Equations of Motion

A dynamical system may be regarded as consisting of \( N \) (a large finite number) particles\(^1\) with masses \( m_1, m_2, \ldots, m_N \) whose positions are determined by \( N \) coordinates \( q_1, q_2, \ldots, q_N \in \mathbb{R}^3 \) and velocities defined as \( \dot{q}_1, \dot{q}_2, \ldots, \dot{q}_N \). The mechanical properties of the system are characterised by two quantities, the kinetic and potential energy. The kinetic energy \( T(\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_N, q_1, q_2, \ldots, q_N, t) \) is a function of velocities in quadratic form:

\[ T = \frac{1}{2} \sum_{i=1}^{N} m_i |\dot{q}_i|^2 \]  

(B.3)

The potential energy \( U(q_1, q_2, \ldots, q_N, t) \) is assumed to be a function of the coordinates \( q_i \) and \( t \). The energy difference

\[ L = T - U \]  

(B.4)

is generally called the kinetic potential or Lagrangian function. Then Hamilton’s principle states that between two instants of time \( t_1 \) and \( t_2 \) the actual motion of the system proceeds in such a way that the integral

\[ J = \int_{t_1}^{t_2} L dt \]  

(B.5)

is stationary with respect to all neighbouring virtual motions which lead from the initial to the final position of the system in the same interval of time. In other words, Hamilton’s principle dictates that the true path of the system make the time integral of the difference between its kinetic energy and potential energy stationary.

Hamilton’s principle leads immediately to Lagrange’s equations of motion. According to (B.1) we have

\[ \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = 0 \quad (i = 1, 2, \ldots, N) \]  

(B.6)

or equivalently

\[ \frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial (T - U)}{\partial q_i} = 0 \quad (i = 1, 2, \ldots, N) \]  

(B.7)

B.3 Euler’s Equations for Variational Problems with Several Independent Variables

Before we deduce Lagrange’s equation of motion for the dynamic deformable models, we need to establish Euler’s equations for variational problems with more than one independent variable, that is, the problem of determining extrema of multiple integrals. Let us

\(^1\)By a particle we mean an object which possesses mass but has no size.
consider the problem of finding the stationary value of the following double integral

\[ I = \iint_G F(x, y, u, u_x, u_y, u_{xx}, u_{xy}, u_{yy}, \ldots, u_{yy \ldots y}) dx dy \] (B.8)

over a given region of integration \( G \) by determining a suitable function \( u(x, y) \) which is continuous, has continuous partial derivatives \( u_x, u_y, u_{xx}, u_{xy}, u_{yy}, \ldots, u_{yy \ldots y} \) up to the \( n \)-th order, and takes on prescribed values on the boundary of \( G \). Just as the variational problems considered in Appendix A lead to ordinary differential equations, this leads to partial differential equations for the functions concerned. Then the following Euler differential equation holds:

\[
[F]_u = F_u - \frac{\partial}{\partial x} F_{u_x} - \frac{\partial}{\partial y} F_{u_y} + \frac{\partial^2}{\partial x^2} F_{u_{xx}} + \frac{\partial^2}{\partial xy} F_{u_{xy}} + \frac{\partial^2}{\partial y^2} F_{u_{yy}} + \cdots + (-1)^n \frac{\partial^n}{\partial y^n} F_{u_{yy \ldots y}} = 0
\] (B.9)

### B.4 Lagrange’s Equations of Motion for Dynamic Deformable Models

From section 9.5 the Lagrangian function for dynamic deformable models is

\[
\mathcal{L}(\mathbf{v}) = \int_0^1 \frac{1}{2} \mu(s) |\dot{\mathbf{v}}|^2 - \frac{1}{2} (\alpha |\mathbf{v}'|^2 + \beta |\mathbf{v}''|^2) - P(\mathbf{v}) ds
\] (B.10)

Let \( L \) represent the integrand in the above equation, and according to (B.9) Lagrange’s equation of motion for dynamic deformable models follows

\[
\frac{\partial L}{\partial \mathbf{v}} - \frac{d}{dt} \left( \frac{\partial L}{\partial \dot{\mathbf{v}}} \right) - \frac{\partial}{\partial s} \left( \frac{\partial L}{\partial \mathbf{v}'} \right) + \frac{\partial}{\partial s^2} \left( \frac{\partial L}{\partial \mathbf{v}''} \right) = 0
\] (B.11)

or after multiplying both side by \(-1\) it becomes

\[
\frac{d}{dt} \left( \frac{\partial L}{\partial \mathbf{v}} \right) - \frac{\partial L}{\partial \mathbf{v}} + \frac{\partial}{\partial s} \left( \frac{\partial L}{\partial \mathbf{v}'} \right) - \frac{\partial}{\partial s^2} \left( \frac{\partial L}{\partial \mathbf{v}''} \right) = 0
\] (B.12)

which is (9.25).
B.4. LAGRANGE’S EQUATIONS OF MOTION FOR DYNAMIC DEFORMABLE MODELS
Bibliography


frequency and envelope-detected signals with applications to medical ultrasound,”

[26] Li Weng, J.M. Reid, P.M. Shankar, K. Soetanto, and X.M. Lu, “Nonuniform phase
distribution in ultrasound speckle analysis—part I: background and experimental
demonstration,” IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency

Mayo Graduate School, USA, 1995.

“Probability density of myocardial ultrasonic backscatter,” in Proceedings of the
1988 Fourteenth Annual Northeast Bioengineering Conference. 1988, pp. 305–308,
IEEE.

[29] S. Miyata and M. Ohta, “A study on the statistical characteristics of speckle in

using marked regularity models,” IEEE Transactions on Ultrasonics, Ferroelectrics

[31] E. Jakeman and P.N. Pusey, “A model for non-Rayleigh sea echo,” IEEE Transac-

IEEE Transactions on Aerospace and Electronic Systems, vol. 27, no. 2, pp. 238–

[33] R.C. Molthen, V.M. Narayanan, P.M. Shankar, J.M. Reid, V. Genis, and L. Vergara-

[34] I.R. Joughin, D.B. Percival, and D.P. Winebrenner, “Maximum likelihood estimation
of K distribution parameters for SAR data,” IEEE Transactions on Geoscience and

non-Rayleigh statistics for the identification of tumors in ultrasonic B-scans of the

[36] V.M. Narayanan, P.M. Shankar, and J.M. Reid, “Non-Rayleigh statistics of ultra-
sonic backscattered signals,” IEEE Transactions on Ultrasonics, Ferroelectrics and


BIBLIOGRAPHY


BIBLIOGRAPHY


[196] Special Issue on Wavelets, vol. 84, no. 4 of Proceedings of the IEEE, April 1996.
BIBLIOGRAPHY


BIBLIOGRAPHY


289


BIBLIOGRAPHY


List of Publications

So far the following publications have resulted from the study towards this thesis:


