

MULTI-SCALE METHODS APPLIED TO SMALL-ANGLE X-RAY  
SCATTERING IMAGES OF BREAST TUMOUR TISSUE

By

Gregory Alan Falzon

A thesis submitted for the degree of  
**Doctor of Philosophy**  
of The University of New England

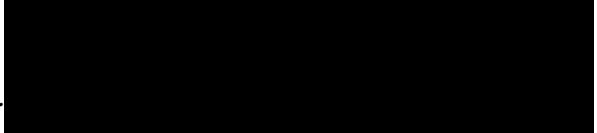
*April 2008*

## **DECLARATION**

---

*I certify that the substance of this thesis has not already been submitted for any degree and is not currently being submitted for any other degree.*

*I certify that to the best of my knowledge, any help received in preparing this thesis, and all sources used, have been acknowledged in this thesis.*

.....  .....

---

# Acknowledgements

First and foremost I would like to take this opportunity to thank my principal supervisor Associate Professor Murison for all of the dedication, hard work and sound guidance that led to the completion of this project. Thank you also to my supervisors Dr.C.Hall and Dr.S.Pearson for arranging funding, leading experiments, collect data and participating in this project.

Thanks to Prof. R.Lewis, Dr. K.Siu, Dr. S. Wilkinson, Dr. A. Round and Dr. S.Sidhu for their joint involvement in the imaging, collection and pre-processing of the data. The technical assistance of Dr. G. Grossmann and Dr. T. Geraki at Daresbury Laboratory is also greatly appreciated. Thank you to Ms. J. Delves for assisting with the drafting of this document.

The access to facilities and the technical instrumentation support provided by CCLRC Daresbury Laboratory Synchrotron Radiation Source is also recognised.

I would like to acknowledge the participation of the medical and nursing staff of the Monash Medical Centre, Nottingham Hospital, the Royal Melbourne Hospital and Southern Health. I also recognise the important role of the patients who donated the tissue samples which enabled the collection of the SAXS data.

Funding for this project was provided by the Australian Government in the form of an Australian Postgraduate Award This project was supported in part by the Australian Research Council's Discovery funding scheme (Project Number DPO662961). Funding was also obtained by the University of New England and the Access to Major Research Facilities Program, which is provided by the Commonwealth of Australia.

Finally, I would like to acknowledge the comments of the examiners who have taken the time to read this thesis and provided useful feedback to help improve the quality of this work.

# Abstract

This thesis develops statistical methods for investigating small angle x-ray scattering (SAXS) images of breast tissue of three different pathologies: normal, benign and malignant. The objective was to thoroughly examine the images and detect those features indicative of malignancy. The tissue sample which is the source of the SAXS image is approximately ten millimetres long and one millimetre wide. In comparison, the structures of interest are on the nanometre scale and therefore we would expect that in practice each sample is a mixture of the different tissue pathologies. This mixture of different pathologies will have a bearing on the overall classification of the sample. Conventional classification models will tend to use data that has been reduced to those components that show the most variation. However in this case, even trivial amounts of feature suggestive of malignancy must be retained as they might be influential in the classification of tissue type.

The mathematical strategy adopted in this thesis relied on a series of transforms along with the resulting interpretation of their coefficients. An adaptive transform that used a range of filter functions was applied to the SAXS images. The coefficients from this transform that had an acceptable probability of misclassification were retained and the others rejected.

Probability density functions were estimated for each group, filter, scale and location index of the retained adaptive image transform coefficients. Where possible a bivariate Gaussian probability density function was used to obtain a parametric estimate, where this was not possible non-parametric estimates were obtained by smoothing bivariate histograms of coefficient magnitude using the Mexican hat contourlet transform. Smoothing was achieved by first modifying the Mexican hat contourlet transform coefficient magnitude according to a generalised extreme value distribution weighting function and then applying the inverse Mexican hat contourlet transform.

The probability density function estimates were then used in Bayes' rule to estimate the probability that the sample belonged to each group for each filter, scale and location indices of the adaptive image transform coefficients. This allowed the identification of those adaptive image transform coefficients most indicative of malignancy.

Three estimates were also found that encoded the overall probability that the sample belonged to the normal, benign or malignant tissue groups. This estimate was found by first averaging over the array of filter, scale and location probability estimates for each group and then estimating univariate probability density functions (of these averages) using the Walsh wavelet packet transform. These probability density functions were then calibrated using an independent *adjustment* data set in order to produce more accurate probability estimates. The overall probability estimates could then be found for future observations by first calculating the naive average probabilities (across filters, scales and locations) for each group and then adjusting these probabilities using the univariate probability density functions. The statistical methodology developed in this thesis was then applied to an *independent* SAXS image data set with good classification results achieved.

# Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>Abstract</b>	<b>iv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation for the Model . . . . .	2
1.1.1 Analysis of large quantities of SAXS images . . . . .	2
1.2 SAXS Imaging in Cancer Research . . . . .	3
1.2.1 SAXS imaging of breast tumours . . . . .	4
1.2.2 Previous image analysis of SAXS data . . . . .	5
1.3 Thesis structure . . . . .	6
1.3.1 Thesis objectives . . . . .	6
1.3.2 Thesis outcomes . . . . .	6
1.3.3 Thesis organisation . . . . .	7
<b>2 Scientific Background</b>	<b>9</b>
2.1 Small-Angle X-ray Scattering . . . . .	10
2.2 Collagen Structure & SAXS . . . . .	12
2.2.1 Collagen structure . . . . .	12
2.2.2 summary . . . . .	14
2.3 Collagen Structure & Breast Cancer . . . . .	15
2.3.1 Overview . . . . .	15
2.4 Chapter summary . . . . .	18
<b>3 Previous SAXS Research on Cancer Diagnosis</b>	<b>19</b>
3.1 Survey of biomedical research using x-ray scattering . . . . .	20

3.2	Diagnostic models related to Collagen Structure . . . . .	21
3.2.1	Model of Lewis <i>et al</i> (2000) . . . . .	22
3.2.2	Round 2006 . . . . .	25
3.2.3	Multivariate analysis: Sidhu <i>et al</i> 2008 . . . . .	31
3.2.4	Summary . . . . .	34
3.3	Diagnostic Models Using Image Analysis . . . . .	35
3.3.1	Data mining: Butler <i>et al</i> 2003 . . . . .	36
3.3.2	Wavelet analysis: Erickson 2005 . . . . .	40
3.3.3	GLMs & wavelet analysis: Falzon <i>et al</i> (2006) . . . . .	47
3.3.4	Section summary . . . . .	50
3.4	Chapter summary . . . . .	50
<b>4</b>	<b>Image Analysis Review</b>	<b>51</b>
4.1	The Sampling Theorem & SAXS Patterns . . . . .	52
4.1.1	The sampling theorem . . . . .	53
4.2	Image Analysis Techniques . . . . .	54
4.2.1	Fourier analysis . . . . .	56
4.2.2	Wavelet analysis . . . . .	59
4.2.3	Fractal analysis . . . . .	62
4.2.4	Binary fractal image analysis . . . . .	63
4.2.5	Gray scale fractal image analysis . . . . .	65
4.2.6	Multi-fractal image analysis . . . . .	65
4.2.7	Surface fractal analysis . . . . .	68
4.2.8	Lacunarity analysis . . . . .	69
4.2.9	Random fields . . . . .	71
4.2.10	Gibbs random fields . . . . .	71
4.2.11	Markov random fields . . . . .	72
4.2.12	Bayes'ian models . . . . .	75
4.2.13	Partial differential equations . . . . .	79
4.2.14	Variational methods . . . . .	82
4.2.15	Mathematical morphology . . . . .	85
4.2.16	Shape analysis . . . . .	89
4.2.17	Fuzzy analysis . . . . .	97

4.3	Summary & Conclusions . . . . .	101
<b>5</b>	<b>Image Transforms</b>	<b>103</b>
5.1	Fourier Transform Family . . . . .	104
5.1.1	Discrete Fourier transform . . . . .	104
5.1.2	Windowed Fourier transform . . . . .	105
5.1.3	Discrete cosine transform . . . . .	106
5.2	The Wavelet Transform Family . . . . .	107
5.2.1	Integral wavelet transform . . . . .	107
5.2.2	Discrete wavelet transform . . . . .	109
5.2.3	Stationary wavelet transform . . . . .	111
5.2.4	The Wavelet packet transform . . . . .	111
5.3	Second generation wavelet-like transforms . . . . .	113
5.3.1	The steerable pyramid . . . . .	113
5.3.2	Curvelets . . . . .	118
5.3.3	Contourlets . . . . .	120
5.4	Adaptive Image Transforms . . . . .	123
5.4.1	Best orthogonal basis . . . . .	123
5.4.2	Matching pursuit . . . . .	124
5.4.3	Basis pursuit . . . . .	126
5.5	Summary & Conclusions . . . . .	127
<b>6</b>	<b>Pre-processing of SAXS images</b>	<b>129</b>
6.1	Technical Specifications . . . . .	129
6.2	Imaging Effects . . . . .	131
6.2.1	Image formation & noise . . . . .	131
6.2.2	Point-spread function . . . . .	131
6.2.3	Beam-stop . . . . .	132
6.2.4	Capillary flare . . . . .	132
6.2.5	Insufficient exposure . . . . .	133
6.2.6	Cosmic ray artifact . . . . .	133
6.2.7	Summary . . . . .	134
6.3	Pre-processing . . . . .	135
6.3.1	Remove wire structure . . . . .	135

6.3.2	Detector flat field response . . . . .	136
6.3.3	Beam intensity normalisation . . . . .	136
6.3.4	Background removal . . . . .	137
6.3.5	Position indexing . . . . .	137
6.4	Summary & Conclusions . . . . .	138
<b>7</b>	<b>Adaptive Transformation of SAXS images</b>	<b>139</b>
7.1	Rationale . . . . .	142
7.2	Adaptive Image Transformation . . . . .	143
7.3	The Library of Transform Functions . . . . .	146
7.3.1	Gamma filter . . . . .	148
7.3.2	Mellin filter . . . . .	155
7.3.3	Chebyshev filter . . . . .	160
7.3.4	Zeta filters . . . . .	165
7.3.5	Witch's Hat filter . . . . .	170
7.4	Filter Function Selection . . . . .	173
7.5	Summary and Conclusion . . . . .	180
<b>8</b>	<b>Classification Model</b>	<b>181</b>
8.1	Modeling Objectives and Challenges . . . . .	185
8.2	Statistical Modeling . . . . .	188
8.2.1	Probability estimates for each coefficient . . . . .	190
8.2.2	XOR estimation . . . . .	192
8.2.3	Parametric densities . . . . .	194
8.2.4	The Mexican hat contourlet transform . . . . .	198
8.2.5	Non-parametric density functions . . . . .	202
8.2.6	Smoothing model . . . . .	210
8.2.7	Contourlet smoothing algorithm . . . . .	216
8.2.8	Threshold policy . . . . .	221
8.2.9	Probability arrays . . . . .	240
8.2.10	Summary . . . . .	241
8.3	Overall Probability Estimates . . . . .	243
8.3.1	Overview . . . . .	243
8.3.2	Naive probability matrix . . . . .	247

8.3.3	Raw histograms . . . . .	248
8.3.4	Walsh wavelet transform . . . . .	249
8.3.5	Wavelet packet smoothing . . . . .	252
8.3.6	Section summary . . . . .	255
8.4	Summary & Conclusion . . . . .	256
<b>9</b>	<b>Application to Breast Cancer Diagnosis</b>	<b>259</b>
9.1	Application of the Adaptive Image Transform . . . . .	261
9.1.1	Data set . . . . .	261
9.1.2	Adaptive image transform . . . . .	262
9.2	SAXS Breast Tissue Classification Model . . . . .	268
9.2.1	Implementation of the XOR model . . . . .	268
9.2.2	Density function estimation . . . . .	271
9.3	Estimating Overall Probability . . . . .	280
9.3.1	Conditional histograms . . . . .	280
9.4	Test Data . . . . .	283
9.5	Summary & Conclusions . . . . .	286
<b>10</b>	<b>Summary, Conclusions, Future Directions</b>	<b>287</b>
10.1	Summary . . . . .	287
10.2	Conclusion . . . . .	288
10.3	Future Directions . . . . .	291
10.3.1	Image analysis . . . . .	291
10.3.2	Statistics . . . . .	292
10.3.3	Applied physics . . . . .	293
<b>Appendices</b>		<b>295</b>

# List of Tables

3.1	Classification model results: Sidhu <i>et al</i> (2008). . . . .	33
3.2	Confidence intervals of mean of the wavelet coefficients features . . . . .	44
8.1	Simulated data results . . . . .	229
8.2	Coefficient absolute maxima . . . . .	234
9.1	SAXS image data sets . . . . .	261
9.2	Bounds on probability of misclassification . . . . .	266
9.3	Proportions of coefficients modelled non-paramterically . . . . .	269
9.4	Classification model assessment . . . . .	284
9.5	Mexican hat wavelet model I . . . . .	285
9.6	Mexican hat wavelet model II . . . . .	285

# List of Figures

2.1	The mechanism of Rayleigh scattering . . . . .	11
2.2	SAXS from oriented collagen fibrils . . . . .	14
2.3	SAXS images of different breast tissue pathologies . . . . .	15
2.4	Electron microscopy images of different breast tissue pathologies . . . . .	16
3.1	The radial integration technique . . . . .	23
3.2	Scatterplots from Round (2006). . . . .	27
3.3	Analysis of SAXS patterns of breast tissue: Sidhu <i>et al</i> (2008). . . . .	32
3.4	Feature extraction: Butler <i>et al</i> (2003). . . . .	36
3.5	Wavelet analysis of the SAXS images: Falzon <i>et al</i> (2006). . . . .	47
4.1	Fourier transform pattern recognition . . . . .	56
4.2	Deterministic and quasi-fractals . . . . .	62
4.3	Box-counting a pathology slide . . . . .	64
4.4	Multi-fractal image analysis . . . . .	66
4.5	Surface fractal image analysis . . . . .	69
4.6	Mammogram image segmentation . . . . .	74
4.7	Bayes'ian image analysis . . . . .	75
4.8	PDE image analysis . . . . .	81
4.9	Variational image inpainting . . . . .	82
4.10	Structuring elements . . . . .	85
4.11	Morphological operations . . . . .	86
4.12	Mathematical morphology in mammography . . . . .	88
4.13	Snake feature extraction . . . . .	91
4.14	Shape analysis of fossils . . . . .	94
4.15	PCA of tooth shape . . . . .	96
4.16	Fuzzy image analysis . . . . .	98

4.17	Fuzzy image segmentation . . . . .	99
5.1	Filters used in the steerable pyramid transform . . . . .	115
6.1	Capillary flare artifact . . . . .	132
6.2	Removal wire structure processing . . . . .	135
6.3	Rat tail collagen SAXS image . . . . .	138
7.1	Analysis strategy . . . . .	139
7.2	Self-similarity of collagen fibril . . . . .	143
7.3	1-D Gamma filter . . . . .	148
7.4	2-D Gamma filter-Fourier space . . . . .	153
7.5	2-D Gamma filter . . . . .	154
7.6	1-D Mellin filter . . . . .	155
7.7	2-D Mellin filter . . . . .	159
7.8	1-D Chebyshev filter . . . . .	160
7.9	2-D Chebyshev filter . . . . .	164
7.10	1-D Zeta filter . . . . .	165
7.11	2-D Zeta filter . . . . .	169
7.12	1-D Witch's Hat filter . . . . .	170
7.13	2-D Witch's Hat filter . . . . .	172
7.14	Schematic diagram to represent $\mathcal{D}^*$ . . . . .	178
8.1	Density estimates of the coefficients of image features . . . . .	195
8.2	Scaling function data extraction . . . . .	205
8.3	Detail functions $\rho_{j,l_j,k_1,k_2}(x, y)$ . . . . .	206
8.4	Schematic diagram to represent $\mathcal{D}^*$ . . . . .	216
8.5	Mapping the $Z_{g_1}$ matrix. . . . .	217
8.6	Maxima of Mexican contourlet coefficients: heteroskedastic noise . . . . .	226
8.7	Marron & Wand (1992) densities. . . . .	233
8.8	Coefficient weighting function. . . . .	235
8.9	Wavelet packet crystals . . . . .	251
8.10	Density function scoring table . . . . .	253
9.1	Witch's Hat adaptive image transform coefficients . . . . .	263
9.2	Probability of misclassification across locations . . . . .	265
9.3	XOR model across scales . . . . .	268
9.4	Density functions for amorphous scatter . . . . .	271

9.5	Raw histogram . . . . .	272
9.6	Mexican hat contourlet transform coefficients . . . . .	273
9.7	Weighting functions . . . . .	274
9.8	Probability density function estimates . . . . .	275
9.9	Individual scale-location probability estimates . . . . .	277
9.10	Density adjustment . . . . .	282

# **Publications arising from this thesis**

## **Journal articles**

Falzon G, Pearson S, Murison R, Hall C, Siu KKW, Evans A, Rogers K, Lewis R 2006 ‘Wavelet-based feature extraction applied to small-angle x-ray scattering patterns from breast tissue: a tool for differentiating between tissue types’, *Physics in Medicine and Biology*<sup>rank A</sup>, **51**(10): 2465-2477.

Falzon G, Pearson S, Murison R, Hall C, Siu KKW, Round A, Schültke E, Kaye AH & Lewis R 2007 ‘Myelin structure is a key difference in the x-ray scattering signature between meningioma, schwannoma and glioblastoma multiforme’, *Physics in Medicine and Biology*<sup>rank A</sup>, **52**(21): 6543-6553.

Falzon G, Murison R & Pearson S 2008 ‘Analysis of collagen fibre shape changes in breast cancer’, *Physics in Medicine and Biology*<sup>rank A</sup> **53**(23): 6641-6652.

Sidhu S, Siu KKW, Falzon G, Nazaretian S, Hart SA, Fox JG, Susil BJ & Lewis RA 2008 ‘X-ray scattering for classifying tissue types with breast disease’, *Medical Physics*<sup>rank A</sup>, **35**(10): 4660-4670, also appearing in *Virtual Journal of Biological Physics Research*, (October 1st 2008 issue).

Sidhu S, Siu KKW, Falzon G, Hart S, Fox J & Lewis RA 2009 ‘Mapping of structural changes in breast tissue disease using x-ray scattering’, *Medical Physics*<sup>rank A</sup>, **36**(7):3211-3217.

## Conference papers

Pearson SJ, Siu KKW, Hall C, Reid C & Falzon G 2006 ‘Small-angle x-ray scattering and second-harmonic generation imaging studies of collagen in invasive carcinoma’, *River-Phys: Australian Institute of Physics 17th National Congress 2006*, Brisbane, Queensland, Australia available at <http://www.aip.org.au/content/congress2006#bmp>.

Siu KKW, Sidhu S, Falzon G, Nazarethian S, Hart SA, Fox JG & Lewis RA 2010 ‘Small Angle X-ray Scattering Signatures for Breast Cancer Detection: Direct Comparison of Synchrotron and Laboratory X-ray Sources’, *World Congress on Medical Physics & Biomedical Engineering*, Sep. 7-12, 2009, Munich, Germany, Springer Berlin Heidelberg, IFMBE Proceedings **25**(II):607-610.



