



Provided by the author(s) and University of Galway in accordance with publisher policies. Please cite the published version when available.

Title	The Development of Ultra Wideband Scanning Techniques for Detection and Classification of Breast Cancer
Author(s)	Conceicao, Raquel
Publication Date	2010-09-30
Item record	http://hdl.handle.net/10379/1821

Downloaded 2024-03-13T08:14:45Z

Some rights reserved. For more information, please see the item record link above.



**The Development of Ultra Wideband Scanning Techniques
for Detection and Classification of Breast Cancer**

A dissertation presented

by

Raquel Cruz da Conceição

to

**The College of Engineering and Informatics
National University of Ireland Galway**

in fulfilment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Electrical & Electronic Engineering

Supervisors

Dr. Edward Jones and Dr. Martin Glavin

National University of Ireland Galway

Galway, Ireland

September 2010

Table of Contents

List of Figures.....	vi
List of Tables	xi
Abstract.....	xiii
Acknowledgments	xiv
Acronyms	xv
1. Introduction	1
1.1. Motivation and Background.....	1
1.2. Contributions.....	4
1.3. Chapter by Chapter Overview.....	7
2. Background	8
2.1. Physiological Background	8
2.1.1. Outline of the Anatomy and Physiology of the Breast	8
2.1.2. Formation of Tumours	9
2.1.3. Breast Cancer	12
2.2. Dielectric Properties	13
2.3. Ultra Wideband Radar Microwave Imaging.....	21
2.3.1. FDTD Modelling Method.....	22
2.3.2. Data-Independent Beamforming.....	23
2.3.3. Data-Adaptive Beamforming.....	27
2.4. Classification Techniques	28
2.4.1. Classification in 2D Systems	28
2.4.2. Classification in 3D Systems	30
2.5. Conclusions	31
3. Tumour Modelling	32

3.1.	Introduction	32
3.2.	Basis of the Gaussian Random Sphere Method	33
3.3.	Tumour Models	34
3.4.	Conclusions	39
4.	Comparison of Planar and Circular Antenna Configurations.....	41
4.1.	Introduction	41
4.2.	Existing Antenna Configurations	42
4.3.	Simulation Procedure and Performance Metrics.....	43
4.3.1.	Numerical Model of the Breast.....	43
4.3.2.	UWB Imaging System	45
4.3.3.	Performance Metrics	47
4.4.	Simulation Results and Discussion	48
4.4.1.	Optimum Number of Antennas for Varying Tumour Size	48
4.4.2.	Optimised Number of Antennas for Varying Tumour Location ..	54
4.5.	Conclusions	62
5.	Classification of Breast Tumours in Homogeneous Breast Tissue	65
5.1.	Motivation and Contributions	65
5.2.	Tissue and Tumour Simulation	66
5.3.	Feature Extraction	68
5.3.1.	Principal Component Analysis	69
5.3.2.	Independent Component Analysis	70
5.3.3.	Discrete Wavelet Transform	71
5.4.	Classification.....	72
5.4.1.	Linear Discriminant Analysis	72
5.4.2.	Quadratic Discriminant Analysis.....	73
5.4.3.	Support Vector Machines	74
5.4.4.	Classification Architectures	76

5.5.	Results and Discussion.....	80
5.5.1.	Comparison of PCA, ICA and DWT	80
5.5.2.	Comparison between LDA and QDA Classifiers	81
5.5.3.	Classification Using SVM	84
5.5.4.	Overall Evaluation of Feature Extraction and Classification	87
5.6.	Conclusions	89
6.	Classification of Breast Tumours in Heterogeneous Breast Tissue	91
6.1.	Motivation and Contributions	91
6.2.	Tissue Modelling.....	92
6.2.1.	Modelling with Fixed Location of Fibroglandular Tissue.....	94
6.2.2.	Modelling with Varying Location of Fibroglandular Tissue.....	95
6.3.	Results and Discussion.....	97
6.3.1.	Classification with Fixed Location of Fibroglandular Tissue	97
6.3.2.	Classification with Varying Location of Fibroglandular Tissue ..	98
6.4.	Conclusions	99
7.	Conclusions and Discussion	102
7.1.	Summary of Work.....	102
7.2.	Conclusions and Discussion.....	103
7.2.1.	Modelling of Tumours	103
7.2.2.	Antenna Configurations.....	104
7.2.3.	Classification of Tumours in a Homogeneous Tissue	105
7.2.4.	Classification of Tumours in Heterogeneous Tissue	106
7.2.5.	Summary of contributions	107
7.3.	Suggestions for Future Work	108
	References	110
	Appendix. Publications.....	123
A.1.	Published Journal Papers	123

A.2. Published Conference Papers	124
--	-----

List of Figures

Figure 2.1. General anatomy of the breast in frontal and sagittal views. Image adapted from [4].	9
Figure 2.2. Simplified sagittal view of the breast with the representations of two of the most common types of breast tumour: in situ and invasive. Image adapted from [4].	13
Figure 2.3. The variation of the relative permittivity (left) and conductivity (right) of normal and malignant tissue between 3 MHz and 3 GHz, as reported by Chaudhary <i>et al.</i> [17]	14
Figure 2.4. The variation of the relative permittivity (left) and conductivity (right) of tumour tissue, surrounding tissue, and peripheral tissue across the frequency band of 0.02 MHz and 100 MHz, as reported by Surowiec <i>et al.</i> [19].	15
Figure 2.5. The variation of the relative permittivity (left) and conductivity (right) of normal and malignant tissue between 50 MHz and 900 MHz, as reported by Joines <i>et al.</i> [18].	16
Figure 2.6. The variation of the relative permittivity (left) and conductivity (right) of normal and malignant tissue between 0.5 GHz and 30 GHz, as reported by Choi <i>et al.</i> [61].	16
Figure 2.7. The relative permittivity (left) and conductivity (right) of normal breast tissue as measured by Lazebnik <i>et al.</i> [26] over the frequency band 0.5 to 20GHz. Group 1 represents 0-30% adipose tissue, group 2 represents 31-84% adipose and group 3 represents 85-100% adipose tissue.	18
Figure 2.8. The median relative permittivity (left) and conductivity (right) Cole-Cole curves for groups 1, 2 and 3 for normal tissue obtained from reduction surgeries and cancer surgeries. The median relative permittivity curve of the dielectric properties of samples that contained at least 30% malignant tissue content is also shown for comparison. All results correspond to the 50 th percentile [27].	19
Figure 2.9. Generic block diagram illustrating the microwave imaging system.	21

Figure 2.10. The planar antenna configuration is shown on the left when the patient is in a supine position, while the circular antenna configuration is shown on the right when the patient is in a prone position. This image is adapted from [29].	24
Figure 3.1. Samples for different Gaussian Random Spheres (GRSs) representing benign tumour models. From top to bottom: (a) smooth and (b) macrolobulated. The GRS models have an average radius size of 5 mm.	35
Figure 3.2. Samples for different Gaussian Random Spheres (GRSs) representing benign tumour models. From top to bottom: (a) smooth and (b) macrolobulated. The GRS models have an average radius size of 10 mm.	36
Figure 3.3. Samples for different Gaussian Random Spheres (GRSs) representing malignant tumour models. From top to bottom: (a) microlobulated, (b) spiculated with 3 spicules, (c) 5 spicules and (d) 10 spicules. The GRS models have an average radius size of 5 mm.	37
Figure 3.4. Samples for different Gaussian Random Spheres (GRSs) representing malignant tumour models. From top to bottom: (a) microlobulated, (b) spiculated with 3 spicules, (c) 5 spicules and (d) 10 spicules. The GRS models have an average radius size of 10 mm.	38
Figure 3.5. Sample of (a) backscattered signals for tumours of identical shape with different sizes and (b) backscattered signals for tumours of different shapes with identical size.	39
Figure 4.1. The supine and prone breast models are shown in (a) and (b), respectively. In the grid area, the white regions represent adipose tissue, while the darker regions represent fibroglandular tissue. The tumour is represented in green. A 2 mm layer of skin limits both models (shown in dark grey), and on top of it the antenna locations are shown in red dots.	44
Figure 4.2. The time-domain (left) and magnitude (right) spectra of the UWB input pulse.	45
Figure 4.3. Simplified block diagram representing the artifact removal algorithm as described in Bond <i>et al.</i> [71].	46
Figure 4.4. Block diagram illustrating the simple delay and sum beamformer for a system with N antennas. The signals recorded in each antenna are time delayed,	

scaled by a scaling factor **k** and then added together, finally the signals are windowed and the energy profile is calculated.47

Figure 4.5. Absolute values for the performance metrics: (a) SCRwb and (b) SCRbb versus the number of antennas, with the presence of a 6 mm tumour with a dielectric variation of the normal breast tissue by $\pm 10\%$. Quadratic curves were fitted to the data. For SCRwb the quadratic equations are given by $y = 0.0003x^2 - 0.0557x + 10.3477$ (circular) and $y = -0.0004x^2 + 0.0437x + 5.5195$ (planar). For SCRbb the quadratic equation are given by $y = -0.0370x^2 + 1.1832x + 0.5994$ (circular) and $y = -0.0068x^2 + 0.2333x + 3.8701$ (planar).49

Figure 4.6. Images of backscattered energy (on a dB scale) for the circular antenna configuration, using 16 antennas. For all images, there is a tumour present with its centre at (3.6,-3.6) cm. The size of the tumours is the following: (a) 2mm, (b) 4mm, (c) 6mm, (d) 8mm, (e) 10mm. Note different intensity scales for the different plots.50

Figure 4.7. Images of backscattered energy (on a dB scale) for the planar antenna configuration, using 16 antennas. For all images, there is a tumour present with its centre at (5,-2.4) cm. The size of the tumours is the following: (a) 2mm, (b) 4 mm, (c) 6 mm, (d) 8 mm, (e) 10 mm. Note different intensity scales for the different plots.51

Figure 4.8. Absolute values for four performance metrics varying the relative permittivity from ± 10 to $\pm 50\%$, in steps of 10%, using the circular and planar configurations with 16 monostatic antennas in an FDTD breast model with the presence of a 6 mm tumour.53

Figure 4.9. The planar and circular antenna configurations are shown on the left and right, respectively. The different locations (0d, 1d and 2d) of the 6 mm tumours are shown in green circles. For the planar configuration the tumour is located at: (5,-2.4), (3,-2.4) and (1,-2.4) cm. For the circular configuration the tumour is located at: (3.6,-3.6), (4.1,-3.95) and (4.6,-4.3) cm.55

Figure 4.10. SCRwb results versus the number of antennas used for both planar and circular configurations. Quadratic equations have been fitted to the data as follows: (a) $y = -0.005x^2 + 0.211x + 9.355$ (circular) and $y = -0.009x^2 + 0.291x + 11.463$ (planar); (b) $y = -0.009x^2 + 0.356x + 8.216$ (circular) and $y =$

$-0.010x^2 + 0.304x + 11.186$ (planar); (c) $y = -0.018x^2 + 0.666x + 6.483$ (circular) and $y = -0.006x^2 + 0.155x + 10.421$ (planar).....56

Figure 4.11. SMR results versus the number of antennas used for both planar and circular configurations. Quadratic equations have been fitted to the data as follows:

(a) $y = -0.047x^2 + 1.472x - 2.327$ (circular) and $y = -0.029x^2 + 0.968x + 4.248$ (planar); (b) $y = -0.025x^2 + 0.906x + 0.792$ (circular) and $y = -0.026x^2 + 0.869x + 3.948$ (planar); (c) $y = -0.015x^2 + 0.728x - 0.473$ (circular) and $y = -0.010x^2 + 0.269x + 4.990$ (planar).....57

Figure 4.12. Images of backscattered energy (on a dB scale) for the planar antenna configuration using 17 antennas. A 6 mm tumour is centred at: (a) (5,-2.4) cm, (b) (3,-2.4) cm, (c) (1,-2.4) cm. Note different intensity scales for the different plots....58

Figure 4.13. Images of backscattered energy (on a dB scale) for the circular antenna configuration using 17 antennas. A 6 mm tumour is centred at: (a) (3.6,-3.6) cm, (b) (4.1,-3.95) cm, (c) (4.6,-4.3) cm. Note different intensity scales for the different plots.....59

Figure 4.14. Effects of dielectric heterogeneity on both planar and circular antenna configuration. Each row corresponds to the metrics: SCRwb, SMR, M_{diff} and FWMH. Each column corresponds to a tumour located at distance 0, d and 2d from the centre of the breast.61

Figure 5.1. Cross-section of the 3D FDTD space lattice partitioned into Total Field (TF), Scattered Field (SF) and UPML regions, for a homogeneous breast model. The target, a spiculated tumour located at the centre of the TF in this example, is illuminated by a pulsed plane wave propagating in the +z direction (represented by a dark line) and backscatter is recorded at the first observer location: (0,0,-74) mm (represented by a filled circle). All four observation points are represented by small circles in the image.....67

Figure 5.2. Classification architecture in which only shape classifications are applied: CS and FS.....77

Figure 5.3. Classification architectures in which a 1-stage coarse size classification is applied before shape classification: CSCS and CSFS.....77

Figure 5.4. Classification architectures in which a 2-stage fine size classification is

applied before shape classification: FSCS and FSFS.	78
Figure 5.5. The classification architectures in which a direct fine size and/or shape classifications are applied: DFS and DFSFS.	79
Figure 5.6. Accuracy of LDA (black line) and QDA (blue line) classifiers versus the number of principal components used. The accuracy results for the coarse size classifier are shown on the left and the accuracy results for the coarse shape classifier are shown on the right. Note that the axes are different for the two graphs.	80
Figure 5.7. Simple block diagram of the DWT decomposition “tree” with the relevant sub-band highlighted.	81
Figure 6.1. Cross-section of the 3D FDTD space lattice partitioned into Total Field (TF), Scattered Field (SF) and UPML regions, for a heterogeneous breast model. In this example, the target, a spiculated tumour located at the centre of the TF, is illuminated by a pulsed plane wave propagating in the +z direction (represented by a dark line) and backscatter is recorded at the first observer location: (0,0,-74) mm (represented by a filled circle). All four observation points are represented by small circles in the image.....	93
Figure 6.2. Samples of different Gaussian Random Spheres representing benign tumours. The (a) smooth and (b) macrolobulated tumour models are represented in Model I and Model II, respectively. Tumours are represented in blue and fibroglandular clusters in black.	95
Figure 6.3. Samples of different Gaussian Random Spheres representing malignant tumours. The (a) microlobulated and (b) spiculated (10 spicules) tumour models are represented in samples of Model III and Model IV, respectively. Tumours are represented in blue and fibroglandular clusters in black.....	95
Figure 6.4. Sample of backscattered signals for smooth tumour with a 2.5 mm radius in a homogeneous breast model (black), in a heterogeneous breast model with fixed location of fibroglandular cluster (Model II in blue) and in a heterogeneous breast model with varying location of two fibroglandular cluster (Model IV in red).	96

List of Tables

Table 2.1. Dielectric properties of female breast tissue at 3.2 GHz, as reported by Campbell and Land [20].....	15
Table 2.2. Average dielectric properties of female breast tissue at 900 MHz measured in vivo using an active microwave imaging system developed by Meaney <i>et al.</i> [31].....	17
Table 4.1. Debye parameters for the FDTD model and dielectric properties of each tissue at the centre frequency of the input pulse.	44
Table 4.2. Summary table with performance metrics results for both antenna configurations with different tumour sizes when using 16 transmitting/receiving antennas.....	52
Table 4.3. Results for both antenna configurations with tumours located at three different distances from the centre of the breast (0, d and 2d) using 17 antennas.....	59
Table 5.1. Accuracy for size classifier, subsequent shape classifiers and overall size-then-shape classifier combining feature extractions (PCA, ICA and DWT) with LDA and QDA classifiers for eight different classification architectures. The best accuracy for each stage of the classification architecture is highlighted.	83
Table 5.2. Accuracy for size and subsequent shape classifiers and overall size-then-shape classifier using SVM binary classifier for six different classification architectures. The best accuracy for each stage of the classification architecture is highlighted.	85
Table 5.3. Summary table for the classification performance with PCA, ICA and DWT as feature extraction methods, and LDA, QDA and SVM classifiers for a range of different classification architectures. The best accuracy for each stage of the classification architecture is highlighted.....	88
Table 5.4. Averaged results for each combination of feature extraction and classification methods, for partial size and shape sub-classifiers. For each classification method, the feature extraction method showing the best average	

accuracy is highlighted.....	88
Table 6.1. Debye parameters for the FDTD model. Parameters are established in: (1) Lazebnik <i>et al.</i> [26-27] and (2) UWCEM website [83].	94
Table 6.2. Classification performance for Model I.	97
Table 6.3. Classification performance for Model II.	98
Table 6.4. Classification performance for Models III and IV (one and two randomly-located heterogeneous clusters, respectively).	99

Abstract

This thesis addresses the development of imaging techniques for the early detection of breast cancer, based on Ultra Wideband (UWB) radar, a promising emerging technology that exploits the dielectric contrast between normal and tumour tissues at microwave frequencies. Of particular interest in this work are issues related to the optimisation of UWB system design, as well as techniques for classification of potential breast tumours into benign and malignant. This is particularly important given the results from recent studies of the dielectric properties of breast and tumour tissue, which have found that strong similarities exist between the dielectric properties of malignant, benign and normal fibroglandular breast tissue. This creates a more challenging imaging scenario and motivates the development of enhanced signal processing techniques for UWB imaging systems.

The issue of antenna configuration for UWB imaging systems is examined in this thesis through a detailed comparison of two antenna configurations (i) planar, and (ii) circular. UWB signals are simulated by means of the FDTD technique, and the backscattered signals recorded from each antenna configuration are processed by means of a beamformer in order to create UWB images of backscattered energy. The performance of each antenna configuration is evaluated by both quantitative methods and visual inspection, for a number of test conditions.

Tumour growth and development patterns are modelled using Gaussian Random Spheres, using four discrete sizes and four different shapes. Feature extraction methods including Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Discrete Wavelet Transform (DWT), are used to extract the most relevant features from the detailed Radar Target Signatures of the tumours, which are then classified with a number of different classification techniques: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Support Vector Machines (SVM). In addition to these techniques, a number of different multi-stage classification architectures are considered. The feature extraction and classification algorithms are evaluated for both homogeneous and heterogeneous breast tissue models, for a range of different tumour sizes and shapes.

Acknowledgments

I would like to express my most sincere appreciation towards both my supervisors, Dr. Edward Jones and Dr. Martin Glavin. Their unconditional support, patience, encouragement, professionalism and kindness throughout the last three years have made this one of the most enjoyable and challenging experiences of my life. I will be forever grateful for their guidance and trust in me towards the achievement of this degree.

I am sincerely grateful to the staff in the Department of Electrical and Electronic Engineering, NUIG, in particular Prof. Gearóid Ó Laighin, Dr. Maeve Duffy, Dr. Fearghal Morgan, Martin Burke, Myles Meehan, Aodh Dalton and Mary Costello.

I would like to thank Science Foundation Ireland (SFI) who supported this research (grant number 07/RFP/ENEF420).

I wish to express my sincere thanks to Dr. Martin O'Halloran for his endless enthusiasm, guidance and advice throughout the project. The combination of his friendship with his professionalism has been of utter importance to me. A word of appreciation goes to my other colleagues and friends in the department, in particular to Dallan, Sean, Diarmaid, Ronan, Ciaran, Brian, Rob, Pat, Lorcan, Myles and Suzi.

I wish to thank my friends in Galway, for welcoming me so warmly into their lives, in particularly my housemates and the people in the 'Tango community'.

I would like to thank my friends in Portugal who have always been there for me throughout this adventure. A special word of appreciation to Ana Filipa, Pedro, Inês, Marta, Mafalda, Ana e Sílvia. I would also like to thank my Erasmus friends, in particular Laura, Caitriona, Nacho, Eleanor, Pierre, Roberto, Silvia and Valerio.

I am sincerely grateful to my boyfriend Fred for being such an inspiration and for his endless support, patience and kindness.

Finally, I would like to show my appreciation for my family's unconditional support and love throughout the years. A special word of gratitude goes to my sister Eva, my parents Conceição and Helderico, grandma, brother-in-law, cousins, aunts and uncles. I would also like to thank my niece Juliana for all the joy she has brought into my life.

Acronyms

APES	– Amplitude And Phase Estimation
CNR	– Complex Natural Resonance
CS	– Coarse-Shape
CSCS	– Coarse-Size-Coarse-Shape
CSFS	– Coarse-Size-Fine-Shape
DAS	– Delay And Sum
DCIS	– Ductal Carcinoma <i>In Situ</i>
DFS	– Direct-Fine-Shape
DFSFS	– Direct-Fine-Size-Fine-Shape
DMAS	– Delay Multiply And Sum
DWT	– Discrete Wavelet Transform
FDTD	– Finite Difference Time Domain
FS	– Fine-Shape
FS-CLEAN	– Fractional-Sequence CLEAN
FSCS	– Fine-Size-Coarse-Shape
FSFS	– Fine-Size-Fine-Shape
FWHM	– Full Width Half Maximum
GRS	– Gaussian Random Sphere
ICA	– Independent Component Analysis
IDAS	– Improved Delay And Sum
IDC	– Invasive Ductal Carcinoma
ILC	– Invasive Lobular Carcinoma
LCIS	– Lobular Carcinoma <i>In Situ</i>
LDA	– Linear Discriminant Analysis
LDB	– Local Discriminant Bases
LLR	– Log-Likelihood-Ratio
MAMI	– Multistatic Adaptive Microwave Imaging
M_{diff}	– Difference between the actual location of the tumour and the location of the peak in the resulting image of backscattered energy

MI – Microwave Imaging
MIMO – Multiple-Input Multiple-Output
MIST – Microwave Imaging via Space Time
MRI – Magnetic Resonance Imaging
PCA – Principal Component Analysis
PET – Positron Emission Tomography
PML – Perfectly Matched Layer
QDA – Quadratic Discriminant Analysis
QF – Quality Factor
RBF – Radial Basis Function
RCB – Robust Capon Beamformer
RTS – Radar Target Signatures
RWCB – Robust Weighted Capon Beamformer
SC – Selection-Combining
SCB – Standard Capon Beamformer
SCRwb – Signal-to-Clutter Ratio within-breast
SCRbb – Signal-to-Clutter Ratio between-breast
SF – Scattered-Field
SMR – Signal-to-Mean Ratio
SVM – Support Vector Machines
SWCNTs - Single-Walled Carbon NanoTubes
TF – Total-Field
TF/SF – Total-Field/Scattered-Field
UPML – Uniaxial Perfectly Matched Layer
UWB – Ultra Wideband

Chapter 1

Introduction

1.1.Motivation and Background

Breast cancer is one of the most common types of cancer. In the United States, breast cancer is second only to lung cancer as the most incident cancer amongst women, affecting 26% of all female cancer patients [1], while it is the most frequent type of cancer amongst Irish women [2]. Approximately 182,000 cases are diagnosed in the United States [1] and over 2,400 cases are diagnosed in Ireland [2-4], every year. In Ireland alone, the statistics show that 1 in 11 women are likely to develop breast cancer at some stage of their lives, with an incidence of 75% over the age of 50 and an incidence of 37% over the age of 65. Currently, over 25,000 women have survived breast cancer in Ireland. However, in the context of the 27 countries in the European Union, Ireland has the second highest incidence rates for breast cancer and ranks fifth place for highest mortality, as reported in [2].

The key factors in improving both survival rates and quality of life for cancer patients are: reliable diagnosis for early detection, early intervention and reliable monitoring. Statistics in Ireland from 1994-1997 have shown that in the period of 5 years after breast cancer treatment, the survival rate of a breast cancer patient diagnosed at an early stage (stage I) is as high as $97\pm4\%$ and at a later stage breast cancer (stage IV) is as low as $18\pm7\%$ [5], emphasising the need for a reliable method for detection of cancer at the earliest possible stage of tumour development.

One of the most common approaches for screening and diagnosis of breast cancer is the use of electronic imaging. Several imaging techniques are being used in the context of breast imaging. Currently, X-Ray mammography, Magnetic Resonance Imaging (MRI) and Breast Ultrasound are the most common. Additionally, other techniques have also been investigated for breast imaging, such as Scintimammography and Positron Emission Tomography (PET).

X-Ray mammography is the current *de facto* screening method for detecting

non-palpable early-stage breast cancer, being a widespread low-cost technique. However this type of examination has some well-recognised limitations in terms of sensitivity and specificity (the proportion of patients correctly identified as having and not having cancer, respectively), presenting inherent health risk associated with ionising radiation, and causing patient discomfort due to breast compression [6-8]. Moreover, there is a high rate of both reported false positive and false negative cases, particularly common amongst younger women who typically have denser breast tissue. The false positive rate for X-Ray mammography can be as high as 75%, however much more worryingly is the high false negative rate of 34%. False positive results mean that the patient is falsely diagnosed with breast cancer, causing unnecessary distress to the patient and placing an unnecessary financial burden on the health service. Conversely, false negative results describe cancers that are only detected in follow-up examinations, and cancers that are interpreted as benign tumours. The occurrence of false negatives can result in cancers only being detected later in a more developed stage, considerably reducing the success of treatment [6-10].

Another popular technique is Ultrasound, which is mainly used in follow-up examinations when suspicious masses are initially detected with X-Ray mammography and also for examination of women with denser breasts [4, 8]. Ultrasound is a low-cost non-ionising technique which uses ultrasonic waves to help classify a breast lump as a cyst or a solid mass. However, it is difficult to detect and characterise smaller tumours using Ultrasound due to a phenomenon known as speckle. Speckle is caused by the interference of echoes from randomly distributed scatterers that are too small to be resolved within the image, degrading both the contrast and spatial resolution. In Kuhl *et al.* [11], the sensitivity and specificity of Ultrasound were found to be 39.5% and 90.5%, respectively.

Magnetic Resonance Imaging (MRI) is well-known for its high contrast and spatial resolution. Furthermore, this imaging modality does not require the use of ionising radiation as its technology is based on magnetic fields acting on the hydrogen atoms of the different tissues of the body. Breast tumours that have induced neighbouring vascularisation are often detected with MRI, however their classification of the tumours as benign or malignant is often difficult. Furthermore, early stage small tumours with insignificant vascularisation are often missed. It has been found that the sensitivity of MRI ranges between 88% and 100% and the

specificity has a high variability between 28% and 100% [8]. In addition, the high cost of MRI means that it is not a viable breast imaging modality for widespread screening.

The two last imaging modalities, Scintimammography (also known as Nuclear Medicine Breast Imaging) and PET are both significantly expensive and require the use of radioactive agents to increase the contrast between tumour and normal tissues [8]. Although Scintimammography successfully detects vascularised tumours, the number of false positive results ranges between 66% and 73%, and the number of false negative results reduces significantly between 2% and 8%. Regarding PET, its sensitivity ranges between 80% and 100% and its specificity ranges between 75% and 85%, as obtained in small scale studies [8]. However, the radioactive agent used in PET is quite expensive on its own and only lasts a limited period of time.

In this context, Microwave Imaging (MI) has been proposed as an appealing alternative to X-Ray mammography, Breast Ultrasound, MRI, Scintimammography and PET as a breast cancer diagnosis method [12-13], as it is potentially low cost, it is non-invasive, does not subject the patients to ionising radiation nor discomfort, offers good resolution and can potentially be used to detect tumours at an early stage of development. It is particularly attractive as an alternative to X-Ray Mammography as an imaging technique for initial screening.

Microwave Imaging exploits the differences in dielectric properties between the constituent tissues of the breast and cancerous masses at microwave frequencies. The dielectric properties of normal and malignant breast tissues have been comprehensively examined for over 20 years across a number of studies. Generally it has been acknowledged that malignant tumours present high values for conductivity and relative permittivity due to higher concentrations of water, which is caused by abnormal vascularisation [13-14] and the fact that malignant tissues retain more fluid in the form of bound water [15-16]. Although, historically, studies indicate that there is a significant dielectric contrast between normal tissue and malignant tissues at microwave frequencies [15-25], recently Lazebnik *et al.* [26-27] have found that, in dielectric terms, only normal adipose tissue and tumour tissue properties differ significantly whereas normal fibroglandular, malignant and benign tumour tissues have rather similar dielectric properties. Therefore examinations performed on women with denser breasts (usually in younger women) are more likely to be inconclusive or to yield false positive or false negative results.

Three different approaches have been proposed in order to image the breast based on the contrasting dielectric properties at microwave frequencies: UWB Radar Imaging, Microwave Tomography and Time-Reversal FDTD methods. The UWB Radar Imaging involves illuminating the breast with a UWB pulse, and recording the resulting backscattered signals. Subsequently, these recorded signals are processed using a beamformer to identify the presence and location of significant dielectric scatterers within the breast [12, 28-29], which potentially indicate the presence of tumours. Microwave Tomography involves a full reconstruction of the dielectric profile of the breast, by solving a forward and inverse scattering problem. These algorithms seek to minimise the difference between measured and calculated electric fields [30-35]. The Time-Reversal Finite-Difference Time Domain (FDTD) approach involves applying time-reversed FDTD equations to all points of the breast grid, with the wave converging at the tumour point [36-38].

In order to overcome the high ratios of false positive and false negative results, a number of other approaches have been examined. For example, a number of researchers have emphasised the need for classification of suspicious regions within UWB images of the breast in order to identify a tumour as malignant or benign [10, 26, 39]. Some approaches have involved the analysis of characteristics inherent to malignant tumours in the breast, such as shape and texture of surface [13, 40-43]. Previously used classification algorithms include those studied by Davis *et al.* [44], Chen *et al.* [45-47] and Teo *et al.* [48].

1.2.Contributions

The significant contributions of this thesis in the area of UWB radar breast cancer imaging are as follows:

- Creation of a large database of 480 benign and malignant 3D tumour models based on Gaussian Random Spheres (GRSs). The database of tumour models includes tumours of four different stages of development for benign and malignant tumours ranging between smooth, macrolobulated, microlobulated and spiculated models, and four different sizes ranging between 2.5, 5, 7.5 and 10 mm in radius. This is one of the largest tumour databases based on GRSs (for classification by means of UWB) found in the literature, and also

the first to discriminate between smooth and macrolobulated GRSs as benign models.

- A comparison of two antenna configurations, planar and circular, is completed using various performance metrics. The optimum number of antennas for both configurations is found for different imaging scenarios in which tumours of different sizes are placed in different locations of the breast and the robustness of the system is tested with increasing variance of dielectric properties.
- Classification of size and shape of breast tumours when immersed in a 3D breast model by means of their Radar Target Signatures (RTS). Classification is pursued in two different imaging scenarios: a homogeneous breast model, and a breast model that accounts for dielectric heterogeneity based on realistic MRI-based models.
- Comparison between different feature extraction methods for breast cancer classification.
- Different classification methods applied to the RTS of the different tumours are compared. Classification architectures, comprising binary or direct classifiers, as well as different combinations of coarse and/or fine size and/or shape classifiers, are tested and compared.

The following journal and conference publications have resulted from the work presented in this thesis:

Journal publications:

- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Comparison of Planar and Circular Antenna Configurations For Breast Cancer Detection Using Microwave Imaging", *Progress In Electromagnetics Research*, Vol. 99, pp. 1-20, 2009 [49].
- **R. C. Conceição**, M. O'Halloran, E. Jones and M. Glavin, "Investigation of Classifiers for Early-Stage Breast Cancer Based on Radar Target Signatures", *Progress In Electromagnetics Research*, Vol. 105, pp. 295-311, 2010 [50].
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Support Vector Machines for the Classification of Early-Stage Breast Cancer Based on Radar Target Signatures", *Progress In Electromagnetics Research B*, Vol. 23, pp.

311-327, 2010 [51].

- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Evaluation of Features and Classifiers for Classification of Early-Stage Breast Cancer", *Journal of Electromagnetic Waves and Applications*, Vol. 25, pp. 1-14, 2011 [52].
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Effects of Dielectric Heterogeneity in the Performance of Breast Tumour Classifiers", *Progress In Electromagnetics Research M*, Vol. 17, pp. 73-86, 2011 [53].
- M. O'Halloran, **R. C. Conceição**, D. Byrne, M. Glavin and E. Jones, "FDTD Modeling of the Breast: A Review", *Progress In Electromagnetics Research B*, Vol. 18, pp. 1-24, 2009 [54] (contribution in Section 4.2. *Modeling the growth patterns and dielectric properties of benign and malignant breast tissue*).
- B. McGinley, M. O'Halloran, **R. C. Conceição**, F. Morgan, M. Glavin and E. Jones, "Spiking Neural Networks for Breast Cancer Classification Using Radar Target Signatures", *Progress In Electromagnetics Research C*, Vol. 17, pp. 79-94, 2010 [55] (contribution in Section 2. *Tumor Modeling* and in Section 4.1. *Principal Component Analysis*).
- M. O'Halloran, B. McGinley, **R. C. Conceição**, F. Morgan, M. Glavin and E. Jones, "Spiking Neural Networks for Breast Cancer Classification in a Dielectrically Heterogeneous Breast", *Progress In Electromagnetics Research*, Vol. 113, pp. 413-428, 2010 [56] (contribution in Section 2. *Modeling* and in Section 4.1.1. *Discrete Wavelet Transform*).

Conference publications:

- **R. C. Conceição**, D. Byrne, M. O'Halloran, M. Glavin and E. Jones, "Classification of Suspicious Regions within Ultrawideband Radar Images of the Breast", *16th IET Irish Signals and Systems Conference, ISSC 2008*, Galway, Ireland, Vol. 1, pp. 60-65, 2008 [57].
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Antenna configurations for Ultra Wide Band Radar Detection of Breast Cancer", in *Proceedings of the SPIE BIOS West*, San José, CA, USA, Vol. 7169, No. 9, pp. [71691M, 12], 2009 [58].

- **R. C. Conceição**, M. O'Halloran, D. Byrne, E. Jones and M. Glavin, "Tumor Classification Using Radar Target Signatures", *Progress In Electromagnetics Research Symposium*, Cambridge, MA, USA, pp. 346-349, 2010 [59].

1.3.Chapter by Chapter Overview

The remainder of the thesis is organised as follows: Chapter 2 describes the background of the work developed in the thesis in more detail, in particular the physiological background related to the breast and tumours, a historical examination of the study of dielectric properties within the breast and of cancer tissues, a brief review of UWB Microwave Imaging (MI) and a first introduction to classification techniques of breast masses in MI. Chapter 3 introduces the modelling of benign and malignant tumours based on GRSs. Chapter 4 describes a comparison between planar and circular antenna configuration evaluated using simple tumour models of different sizes in different locations of the breast by means of a number of performance metrics. Chapter 5 details the classification of tumours by comparing different feature extraction methods which allow the analysis of the RTS of tumours and different classification methods for a homogeneous breast model, while Chapter 6 addresses the same issue in additional breast models that attempt to model heterogeneity within the breast. Finally the overall conclusions and suggestions for future work are presented in Chapter 7.

Chapter 2

Background

In this chapter the background information necessary for the understanding and design of the simulation procedures and consequent results, is addressed. First the physiological background is discussed in Sub-section 2.1, followed by the review of dielectric properties research in Sub-section 2.2; the algorithms for UWB Radar Microwave Imaging are described in Sub-section 2.3, finally, in Sub-section 2.4 the breast classification techniques developed in UWB imaging are examined.

2.1. Physiological Background

2.1.1. Outline of the Anatomy and Physiology of the Breast

The breast is mainly composed of three types of tissue: breast fat (or adipose tissue), glandular tissue and connective tissue (fibrous strands called Cooper's ligament). The proportions of these main types of tissue may vary from person to person [10, 26, 39] and the amount of water, fat and fibroglandular tissue may also vary due to normal hormonal changes in different stages of menstruation, pregnancy, lactation or menopause [10, 60]. For the purpose of bioelectrical studies, the anatomy of the breast can be simplified and presented as follows:

- Below the skin there is an adipose tissue layer which consists of vesicular cells filled with fat which are collected into lobules and then separated by Cooper's ligament.
- The innermost tissue of the breast consists of mammary glands (lobules that produce milk). Each breast has about 15 to 20 sections termed lobes with many smaller sections of mammary glands, which are arranged in a circular fashion. Each section is terminated by thin tubes, called lactiferous ducts, which connect to a reservoir (also called ampulla) and ultimately connect to

the nipple. These lobes and ducts are also surrounded by Cooper's ligament.

- Cooper's ligament has the function of maintaining the inner structure of the breast and supporting the tissue attached to the chest wall. The breast is separated from the pectoralis major muscle by the retromammary fat [24].

The anatomy of a healthy breast in both frontal and sagittal views is shown in Figure 2.1. It should be noted that although lymph nodes are not constituents of the breast *per se*, they are represented in this figure as breast cancer can be diagnosed through detection of metastasised tumour cells particularly in the axillary lymph nodes, where approximately 50% of breast cancer occur [28, 61].

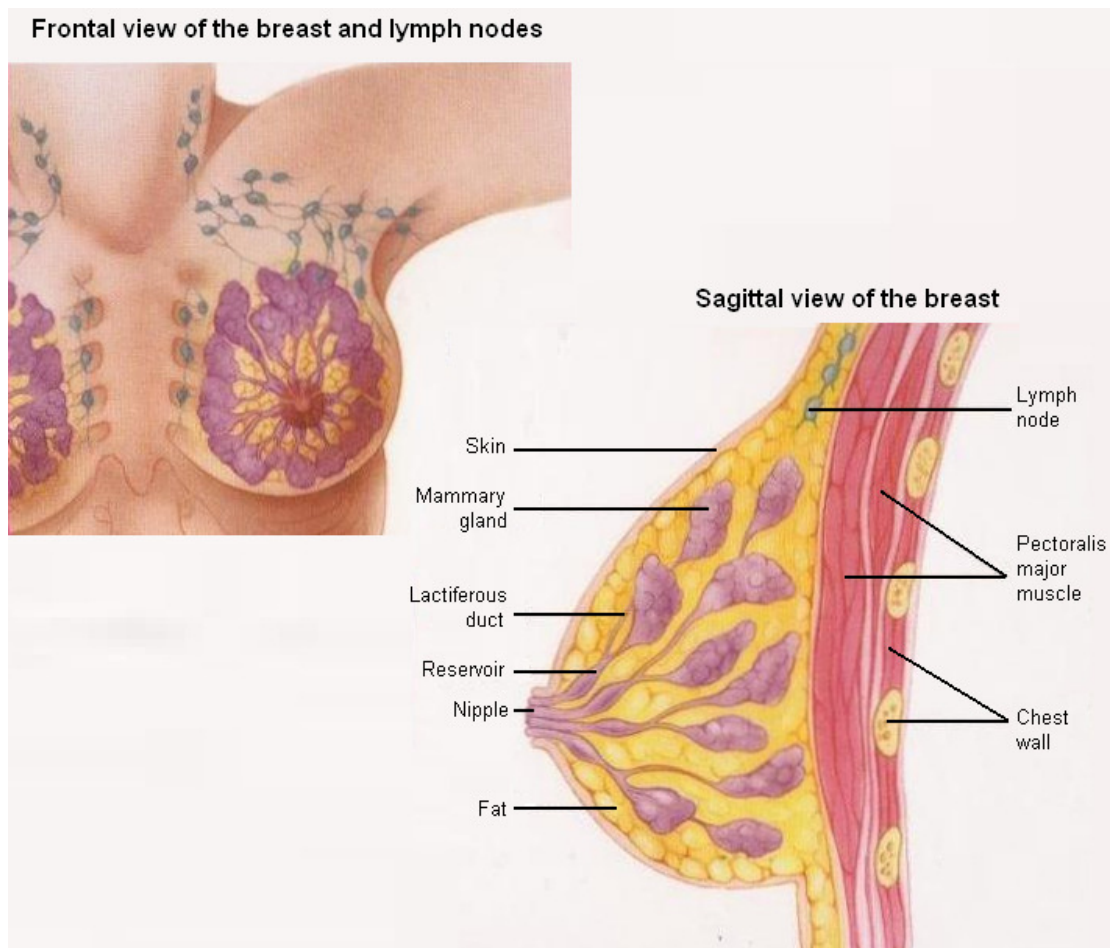


Figure 2.1. General anatomy of the breast in frontal and sagittal views. Image adapted from [4].

2.1.2. Formation of Tumours

Tumours are defined by a growth of undifferentiated (or unspecialised) cells which form a lump. Usually the immune system is capable of destroying the

undifferentiated cells which can lead to the formation of a tumour through a process called apoptosis – cell self-destruction. However, if too many mutations occur in cells at the same time, the immune system may not be able to respond appropriately, and masses of tumour cells will be formed [62].

A carcinogen is the generic name for something that induces the mutation of cells, leading to the formation of tumour cells [62]. Generally, there are two non-exclusive possible origins for tumour cells. The first is the existence of oncogenes, which are the genes responsible for the proliferation of cells, and the second is the repression of the genes that usually restrain cell proliferation and allow apoptosis in order to maintain a controlled overall growth of cells [62-63].

The way proliferation of tumour cells occurs, i.e. the tumour growth, may indicate whether a tumour is benign or malignant. For benign tumours, the growth is controlled and will only be dangerous if nearby organs are pushed and compressed or if tumours either grow inside the skull or release unwanted hormones. Conversely, for malignant tumours, i.e. cancer, the growth is uncontrolled due to a high rate of replication and usually spreads to other parts of the body by metastases and destroys surrounding healthy tissues [62].

The tumour cell suffers several changes in terms of cell surface, the state of water, viscosity, pH, growth regulation, the loss of contact inhibition, the cytoskeleton, temperature, membrane transport, and several other factors [10, 64]. Some of these changes will affect dielectric properties directly, so these will be studied in more detail.

The grade of malignancy of the tumour can be determined by pathologically analysing how premature the cells are within the tumour. The less mature are the tumour cells, the older and more widespread the malignant tumour is likely to be, and therefore the lower the chances of a successful treatment. The different grades of development at which cells can be found is referred to as differentiation [62].

The cytoskeleton of tumour cells becomes disorganised due to the decrease and disorganisation of microfilaments and microtubules [64], causing the original shape of the cell to be lost (becoming more round) and the process of mitosis (cell replication) to become chaotic leading to both an uncontrolled growth of tissues and the loss of genetic information. Because the surface of the cells changes, the membrane permeability is altered and the regular osmosis process is also affected, causing the tumour tissues to retain more fluid than normal cells. Cancer cells retain

more fluid in the form of bound water.

Additionally, cancer cells are not contact-inhibited, which means that huge masses of cancer cells grow over each other and are piled up on one another forming multiple layers, and are able to coexist in high concentrations. Furthermore, due to the large growth of cells in tumours, particularly in malignant tumours, networks of capillaries are created in order to nourish the newly formed cells [13]. In [65], it is noted that neoangiogenesis is induced by tumours with a dimension of at least 3mm. As the size of tumours becomes larger, these networks of capillaries may be developed into tiny veins and even arteries that will connect to major blood supply vessels [13], therefore the study of the level of vascularisation near a tumour is of importance to the characterisation of the grade of malignancy of a tumour.

The increase of water within cancerous tissue is responsible for the high scattering in microwave imaging. The increase of sodium and water, particularly in bound water, within the tumour cell induce greater values of conductivity and relative permittivity in tumour tissues [15-16, 39].

Another feature that may help detect the presence of malignant tumours is the existence of calcifications. However these are only formed when severe necrosis has occurred, i.e. disorderly apoptosis, resulting in groups of dead cells which are not naturally absorbed by the organism [39].

Furthermore, other characteristics inherent to benign and malignant tumours have proven to be useful in terms of classification for different imaging modalities. Such characteristics are based on size, shape, margins, surface texture, depth, localisation and packing density [13, 24, 41, 44, 65]. Features of a tumour that may be of particular benefit for classification in the context of MI are the shape and texture of the tumour surface. Malignant tumours usually present the following characteristics:

- Irregular, ill-defined and asymmetric shapes;
- Blurred boundaries (lack of sharpness);
- Rough and complex surfaces with spicules or microlobules;
- Non-uniform permittivity variations;
- Distortion in the architecture of the breast;
- Irregular increase of tissue density (due to masses and calcifications).

Conversely, benign tumours tend to have the following characteristics:

- Spherical, oval or at least present well-circumscribed contours;
- Compact;
- Smooth surfaces [13, 40-43].

2.1.3. Breast Cancer

Breast cancers are more likely to develop in older women, similarly to what happens with other types of cancer [66]. This is due to the fact that cells have to undergo multiple genetic alterations before a cell becomes malignant [63]. There is also a higher incidence of breast cancer if previous family generations suffered from the same disease (family history) and also if the patient previously developed breast or any other type of cancer [66].

Mainly, there are two most common types of cancer: *in situ* (or non-invasive) and invasive. *In situ* cancers are those in which cancer cells remain within the basement membrane of the lobules and the draining lactiferous duct. Invasive cancers are those in which there is dissemination of cancer cells outside the basement membrane of the ducts and lobules into the surrounding adjacent normal tissue [67].

Some of the most frequently-occurring breast cancers are as follows:

- The Ductal Carcinoma *In Situ* (DCIS) is a type of cancer in which cancerous cells are inside some of the ducts, but have not spread to other regions of the breast or body.
- The Lobular Carcinoma *In Situ* (LCIS) is not a type of cancer *per se*, but in presence of this disease there are high chances of developing cancer. LCIS is characterised by changes in the cells within the breast lobes.
- Invasive Ductal Carcinoma (IDC) is the most common type of breast cancer (70 to 80% of breast cancer cases), and occurs in the cells that line the ducts of the breasts.
- Invasive Lobular Carcinoma (ILC) represents about 10% of breast cancer cases and occurs in the cells that line the lobules of the breast.
- Inflammatory breast cancer is a rare type of breast cancer (1 or 2% of the cases) and induces an inflammation in the breast tissues caused by cancer cells which block the smallest lymph channels in the breast.
- Paget's disease is also rare (1 to 2% of the cases) and is manifested on the

breast skin, usually the nipple or areola (skin area surrounding the nipple), and statistics show that in approximately 90% of the cases this type of cancer is invasive [66].

The most common types of breast tumours, which are also the focus of this thesis, are shown in a simplified sagittal view of the breast in Figure 2.2.

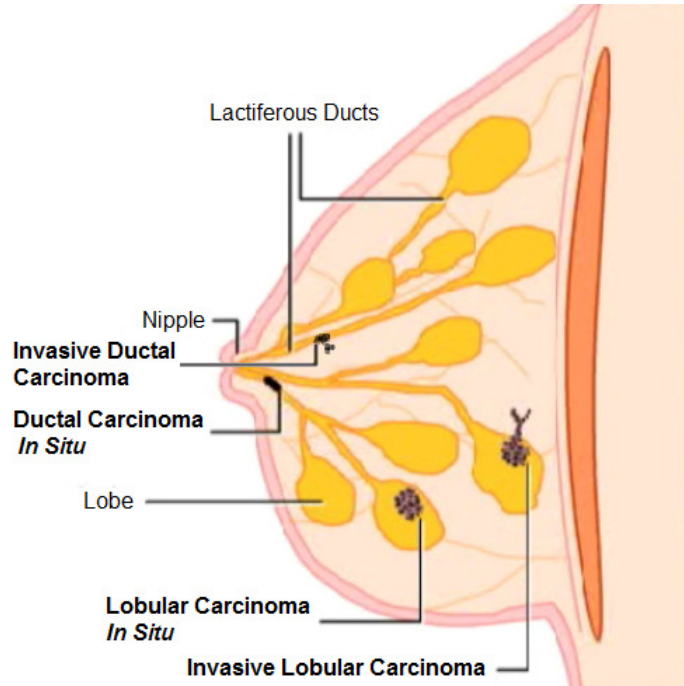


Figure 2.2. Simplified sagittal view of the breast with the representations of two of the most common types of breast tumour: *in situ* and invasive. Image adapted from [4].

2.2.Dielectric Properties

The physiological and biochemical factors associated with breast cancer have been discussed in Sub-sections 2.1.2 and 2.1.3. From a microwave imaging perspective, these factors lead to changes in dielectric properties in the tissue.

The dielectric properties, conductivity and relative permittivity, determine the attenuation of a signal through a medium and the reflections caused by a medium, permitting the differentiation between different types of tissue within the breast at microwave frequencies. Several historical studies have been completed examining the *in vivo* and *ex vivo* dielectric properties of normal and malignant breast tissues in particular, and these are examined in detail in the following paragraphs.

In 1984, Chaudhary *et al.* [17] first examined *ex vivo* specimens of breast tissue removed during cancer surgeries. A significant dielectric contrast between normal

and malignant tissues was found across the frequency range of 3MHz to 3GHz, at 25°C. Chaudhary concluded that significant differences existed in the dielectric properties of normal and malignant tissues of the human female breast, with the greatest dielectric difference occurring at frequencies below 100 MHz. The contrast ratio found for relative permittivity and conductivity was 4.7:1 and 5:1, respectively. Figure 2.3 shows the variation of the dielectric properties of normal and malignant tissue with frequency reported in this paper.

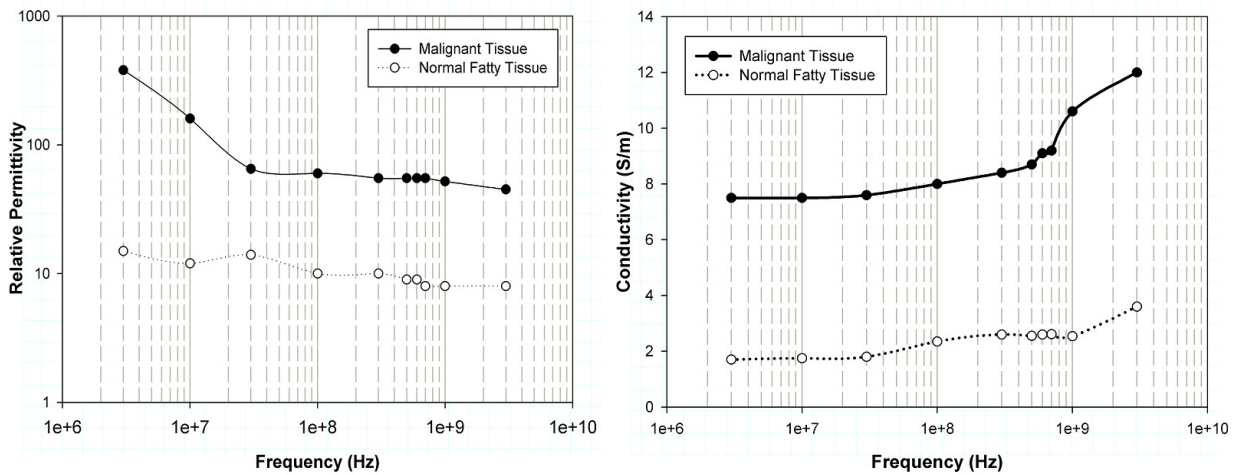


Figure 2.3. The variation of the relative permittivity (left) and conductivity (right) of normal and malignant tissue between 3 MHz and 3 GHz, as reported by Chaudhary *et al.* [17]

In a later study in 1988, Surowiec *et al.* [19] measured the relative permittivity and conductivity of infiltrating breast carcinoma, the surrounding tissue and the peripheral tissue at frequencies between 20 KHz and 100 MHz. The *ex vivo* samples were taken from a population of seven patients and stored in physiological saline. Three measurements were made in three locations: the central part of the tumour, the tissue directly surrounding the tumour and the peripheral tissue approximately 2 cm away from the centre of the tumour. Their results, shown in Figure 2.4, may suggest that there are increased dielectric properties even at the edge of the tumour due to tumour cell proliferation, and that smaller tumours may still be detected using a UWB radar.

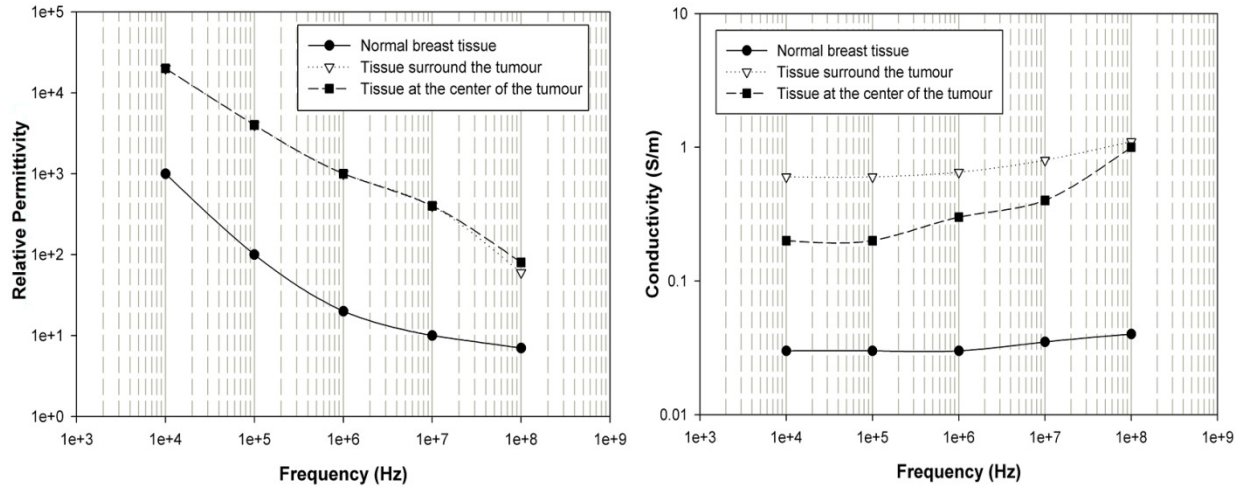


Figure 2.4. The variation of the relative permittivity (left) and conductivity (right) of tumour tissue, surrounding tissue, and peripheral tissue across the frequency band of 0.02 MHz and 100 MHz, as reported by Surowiec *et al.* [19].

In 1992, Campbell and Land [20] provided detailed information for microwave thermography applications on the dielectric properties of *ex vivo* female breast tissue at 3.2 GHz. In this study a resonant cavity technique was used to measure the dielectric properties in four different types of tissue: fat tissue, normal tissue, benign breast tumour and malignant breast tumour. Their results are shown in Table 2.1. They found that dielectric properties overlap for benign and malignant tumour tissues and also observed a much greater range of dielectric properties for normal tissue, suggesting that normal tissue and both benign and malignant tissue may be difficult to differentiate solely based on their dielectric properties.

Table 2.1. Dielectric properties of female breast tissue at 3.2 GHz, as reported by Campbell and Land [20].

Tissue Type	Relative Permittivity	Conductivity (Sm^{-1})	Water content (%)
Fat tissue	2.8-7.6	0.54-2.9	11-31
Normal tissue	9.8-46	3.7-34	41-76
Benign breast tumour	15-67	7-49	62-84
Malignant breast tumour	9-59	2-34	66-79

In 1994, Joines *et al.* [18] measured *ex vivo* samples at frequencies between 50 to 900 MHz, a range typically used for microwave-induced hyperthermia. Taken from 12 patients, tissue samples were analysed and results showed significant differences between normal and malignant tissues for the mammary gland, with a difference ratio of 6.4:1 and 3.8:1 for the relative permittivity and conductivity,

respectively, which is in general agreement with the results reported by Chaudhary *et al.* [17]. Their results are plotted in Figure 2.5.

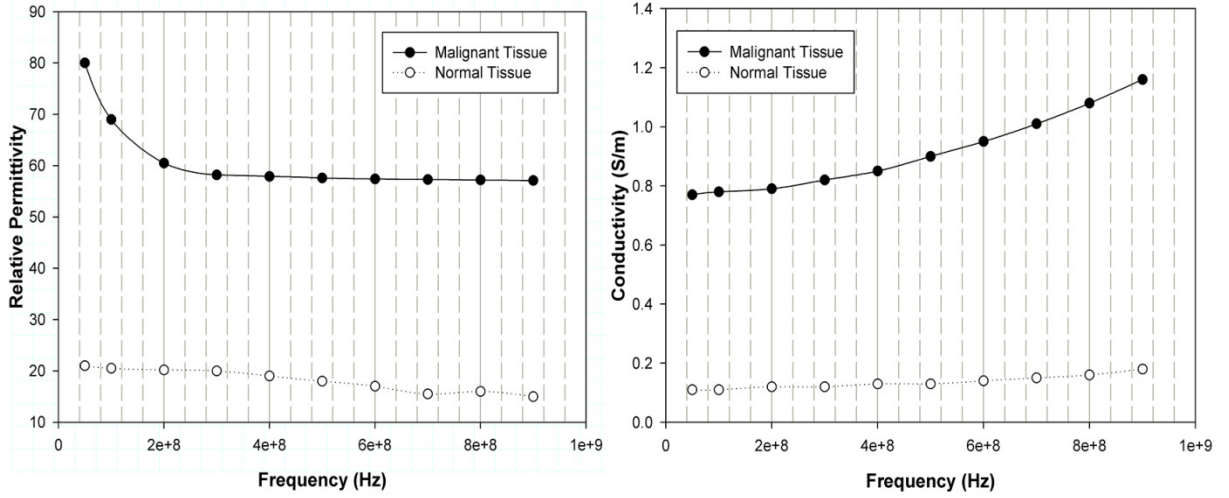


Figure 2.5. The variation of the relative permittivity (left) and conductivity (right) of normal and malignant tissue between 50 MHz and 900 MHz, as reported by Joines *et al.* [18].

Also in 1994, Choi *et al.* [61] examined the dielectric properties of metastasised lymph nodes and normal lymph nodes, along with the dielectric properties of breast cancer tissue in a frequency ranging between 0.5 and 30 GHz. The results are shown in Figure 2.6, and it is observed that both metastasised lymph nodes and breast cancer tissue differ significantly from normal lymph nodes.

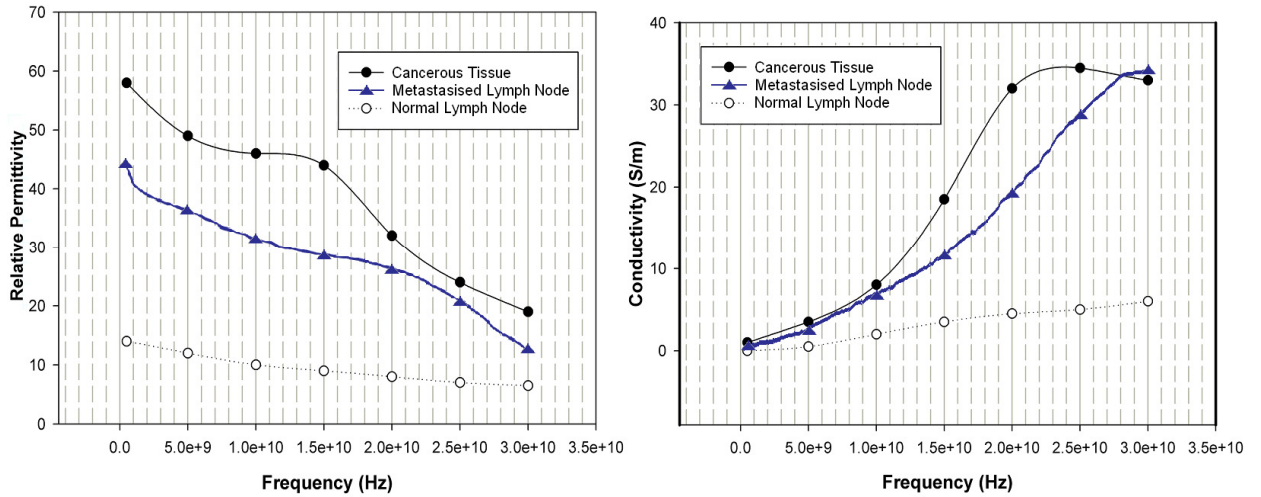


Figure 2.6. The variation of the relative permittivity (left) and conductivity (right) of normal and malignant tissue between 0.5 GHz and 30 GHz, as reported by Choi *et al.* [61].

Meaney *et al.* [31] performed the first clinical analysis *in vivo* using a prototype microwave imaging system, in 2000. A 16 element monopole antenna array was used in a tomographic microwave imaging system between 300 and 1000 MHz.

Results at 900 MHz are shown in Table 2.2 and it can be observed that the average relative permittivity value is significantly higher, approximately between 31 and 36, than that published in Joines *et al.*'s study [18]. No malignant tissue was examined in the course of this study and so a direct comparison cannot be made to the previous *ex vivo* studies.

Table 2.2. Average dielectric properties of female breast tissue at 900 MHz measured in vivo using an active microwave imaging system developed by Meaney *et al.* [31].

Patient	Age	Average relative permittivity (%)	Average conductivity (Sm ⁻¹)
1	76	17.22±11.21	0.5892±0.3547
2	57	31.14±4.35	0.6902±0.3650
3	52	36.44±6.24	0.6869±0.3156
4	49	35.43±3.93	0.5943±0.3841
5	48	30.85±7.22	0.6350±0.3550

More recently, Lazebnik *et al.* [26-27] completed one of the most comprehensive studies to date on dielectric properties of the breast. The first study [26] focused on the dielectric properties of normal tissue, and the second study [27] focused on the dielectric contrast between normal, benign and malignant breast tissues. On both studies, the data were mapped to Cole-Cole representations in order to assist on the measurements of the dielectric properties. Hoping to improve on many of the apparent weaknesses of previous studies, such as small patient sample sizes and gaps in the frequency bands examined, Lazebnik analysed histopathologically a large pool of freshly excised breast tissue from patients and divided normal tissue samples into 3 groups, distinguishing each by the percentage of adipose, glandular and fibroconnective tissue contained in the sample before obtaining the values for the dielectric properties. The three groups were defined as follows:

- Group 1 contains all samples with 0-30% adipose tissue;
- Group 2 contains all samples with 31-84% adipose tissue;
- Group 3 contains all samples with 85-100% adipose tissue.

The main findings in their first study [26] were that breasts with high adipose and low fibroglandular contents presented lower average dielectric properties, whereas breasts with low adipose and high fibroglandular tissues presented higher dielectric properties, which suggested that a wide range of dielectric properties is possible within healthy breasts. Results are summarised in Figure 2.7.

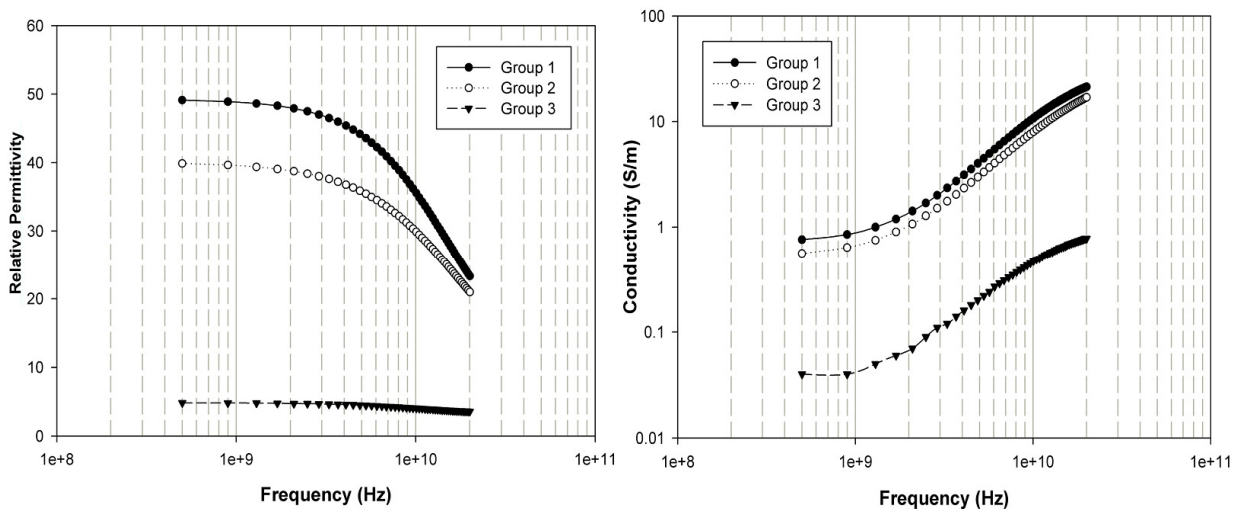


Figure 2.7. The relative permittivity (left) and conductivity (right) of normal breast tissue as measured by Lazebnik *et al.* [26] over the frequency band 0.5 to 20GHz. Group 1 represents 0-30% adipose tissue, group 2 represents 31-84% adipose and group 3 represents 85-100% adipose tissue.

By comparing these results to previous dielectric studies, Lazebnik's main conclusions were as follows:

- The dielectric properties found for normal tissue in the samples of Group 3 (the highest in adipose content) were lower than any previous studies.
- The dielectric properties found for normal tissue in the samples of Group 1 (the highest in fibroglandular content and lowest in adipose content) were higher than any previous studies.
- The dielectric data spanned a much greater range of values than those reported in previous studies, with an exception to Campbell and Land's study [20].

Overall, Lazebnik attributed these differences to the large heterogeneity in normal breast tissue, as previously noted in [20], and acknowledged the relation found between the content of tissues within the breast – more or less adipose – and the measured dielectric properties – lower or higher, respectively.

In their second study, Lazebnik *et al.* [27] further addressed the differences between normal, benign and malignant tumours across a frequency range of 0.5 to 20 GHz. Normal breast tissue included adipose, glandular and fibroconnective tissues; benign tumour tissue included fibroadenoma and cysts; and, finally, malignant tumour included ductal and lobular carcinomas (IDC, DCIS, ILC and LCIS). The results are shown in Figure 2.8.

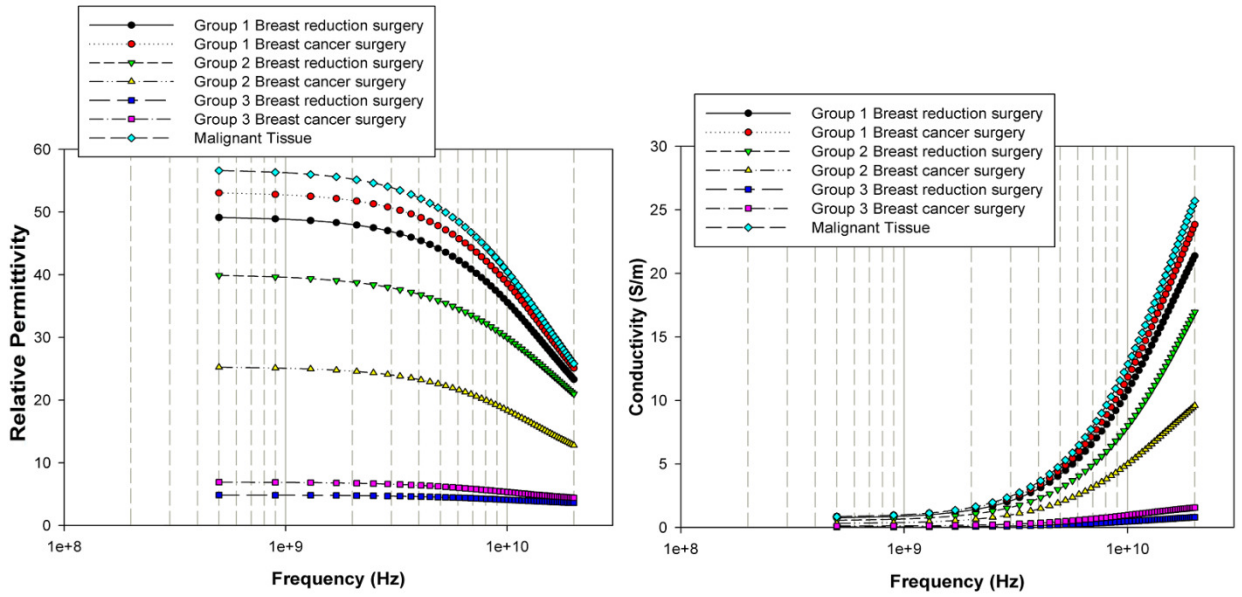


Figure 2.8. The median relative permittivity (left) and conductivity (right) Cole-Cole curves for groups 1, 2 and 3 for normal tissue obtained from reduction surgeries and cancer surgeries. The median relative permittivity curve of the dielectric properties of samples that contained at least 30% malignant tissue content is also shown for comparison. All results correspond to the 50th percentile [27].

Firstly, Lazebnik *et al.* [27] observed that measured dielectric values for malignant tissue were in general agreement with the studies by Chaudhary *et al.* [17], Surowiec *et al.* [19] and Joines *et al.* [18]. Furthermore, Lazebnik *et al.* justified the differences between the curves for group 2 with an experimental error due to the comparatively small sample size used in the cancer surgery study compared to the breast reduction surgery study, which varied from 16 to 84 samples. Also, it was acknowledged in [27] that dielectric properties for normal tissues obtained through breast cancer surgery were lower than those obtained in breast reduction surgery and noted that this is due to the fact that tumours usually develop in glandular tissue and consequently the non-affected tissues removed had comparatively higher content of adipose.

Finally, by adjusting for the content of adipose within the samples Lazebnik *et al.* found that there only existed a 10% difference between the conductivity of normal tissue and malignant tissue, and an approximate 8% difference in relative permittivity at 5 GHz. However, by adjusting for the content of both adipose and fibroconnective tissues within the samples, they found no statistical difference between normal fibroglandular and malignant tumour tissues in the breast. The high

dielectric properties of benign/fibroglandular tissue overlap those of malignant tissue within the breast, and consequently materialise as ‘false positive’ results and create a much more difficult imaging scenario than previously believed.

Overall, the studies in [26-27] greatly added to the knowledge of breast tissue dielectric properties by significantly increasing the population size, by having separate analysis depending on the proportion of different types of tissues within the breast, and by characterising the tissues across a wide frequency band between 0.5 and 20 GHz. In [26-27], a very detailed database of dielectric properties based on Cole-Cole parameters was established for each tissue type, which is crucial to accurately develop a numerical breast phantom, later described in this chapter.

Furthermore, in 2009, Halter *et al.* [68] presented the initial results from a clinical study with a smaller number of patients in which estimations of dielectric properties of malignant breast tissue were obtained in three different scenarios. The dielectric properties were taken (i) by estimation means of Electrical Impedance Tomography imaging, (ii) by direct measurement *in vivo* breast cancer with both Electrical Impedance Spectroscopy (EIS) and Microwave Impedance Spectroscopy (MIS) probes, and finally (iii) by direct measurement in *ex vivo* breast cancer specimens with both EIS and MIS probes. While some of the limitations of this particular study may be due to the small number of patients and the lack of simultaneous measurements for scenarios (i), (ii) and (iii) for each considered lesion, it was noted in this study that there may also be some limitations related to *ex vivo* measurements such as those in Lazebnik *et al.* [26-27].

In [68], it was observed that the dielectric properties for normal breast tissue estimated in scenario (i) agreed with the findings from previous studies, such as [26]. Furthermore, the estimations of dielectric properties in scenario (iii) also agreed well with previous *ex vivo* studies such as [19-20, 25, 27]. However, Haemmerich *et al.* [69] observed a change in some dielectric parameters after excising tissue and attributed those changes to variations of temperature, tissue dehydration and ischemic effects. Moreover, they also noted that those changes may occur within seconds after extraction of tissues and may stabilize for hours thereafter. Finally, for scenario (ii) in Halter’s study, they observed that values for conductivity and relative permittivity of breast cancer were significantly higher than those estimated in scenarios (i) or (iii).

2.3. Ultra Wideband Radar Microwave Imaging

Ultra Wideband Radar Microwave Imaging has been extensively investigated in the context of breast cancer detection for over 10 years [12, 29, 44, 70-73]. Typically, the breast is illuminated with a UWB pulse from antennas placed at a number of different locations surrounding the breast and the resulting backscattered wave signals are recorded at the same antennas, sequentially. In a monostatic system, only one antenna transmits, and the reflected signals are received at the same antenna. In a multistatic system, for each transmitting antenna, all antennas in the system are simultaneously receivers. A microwave beamforming algorithm is applied to the recorded signals so that the backscattered signals can be spatially focused in order to create an image of the dielectric scatterers within the breast, which represent potential tumours. A block diagram of a typical UWB microwave imaging system is shown in Figure 2.9.

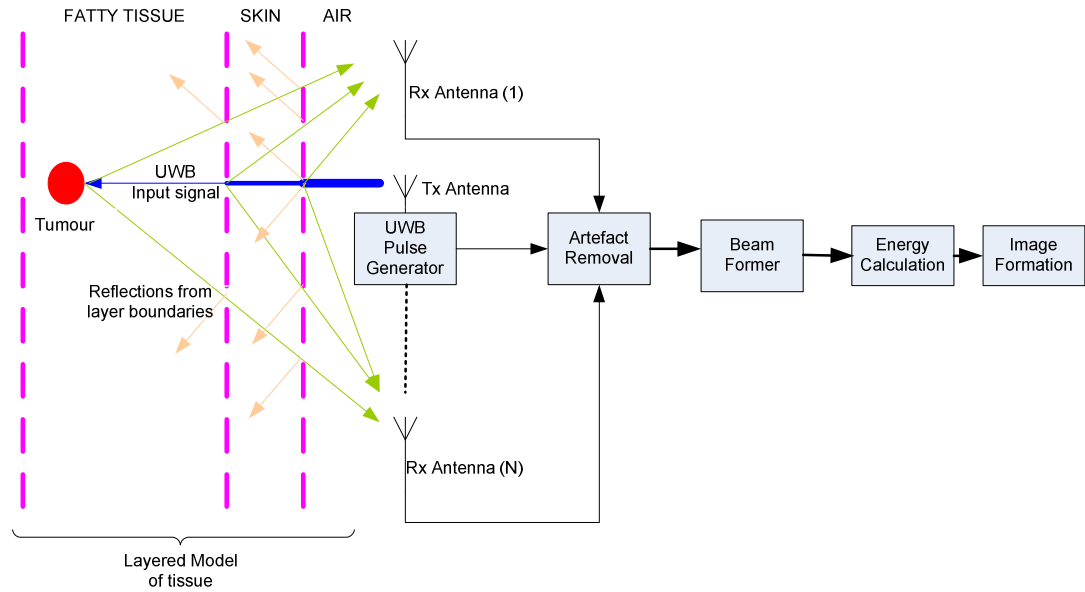


Figure 2.9. Generic block diagram illustrating the microwave imaging system.

To simulate and test UWB imaging algorithms, FDTD modelling is commonly used for generating anatomically and dielectrically accurate breast models. The FDTD method is briefly addressed in Sub-section 2.3.1. In Sub-sections 2.3.2 and 2.3.3 two general types of beamforming are detailed: data-independent beamforming and data-adaptive beamforming, respectively.

2.3.1.FDTD Modelling Method

The Finite Difference Time Domain (FDTD) modelling method is a numerical method commonly used for modelling the propagation of electromagnetic signals in biological tissue, first introduced by Yee [74]. The method is described in detail in the work by Taflove and Hagness [75]. As presented in [75], the main advantages of the FDTD method are as follows:

1. FDTD does not use linear algebra and therefore it allows for FDTD models to have an infinite number of electromagnetic unknowns.
2. FDTD is accurate and robust and sources of errors are well understood.
3. FDTD treats impulse behaviour naturally since it can directly calculate the impulse response of an electromagnetic system, given that FDTD is a time domain method.
4. FDTD treats nonlinear behaviour naturally as it is based on time domain technique.
5. FDTD is a systematic approach since any structure under analysis can be reduced to the design of an appropriate FDTD grid.
6. FDTD has increased capabilities as computer memory capacities are increasing rapidly and this will allow further fine discretisation of space (and time consequently).
7. FDTD profits from increasingly advanced computational visualisation capabilities since the progression of electromagnetic signals can be monitored over stipulated periods of time.

The FDTD method is based on the discretisation of the Maxwell's equations in free space for the electric and magnetic fields, presented in (2.1) and (2.2), respectively.

$$\frac{dE}{dt} = \frac{1}{\epsilon_0} \nabla \times H \quad (2.1)$$

$$\frac{dH}{dt} = \frac{1}{\mu_0} \nabla \times E \quad (2.2)$$

in which ϵ_0 and μ_0 are the permittivity and permeability of free space, respectively. For the discretisation of the above equations, the time and space resolutions, dt and dx , need to be defined. The space resolution is initially set and then the time resolution is calculated by the following equation:

$$dt = \frac{dx}{c_0} \quad (2.3)$$

in which c_0 is the speed of light in free space.

Particularly for the work developed in this thesis, the frequency-dependent nature of biological tissue is incorporated into the FDTD numerical model of the breast through a Debye formulation, defined by Luebbers *et al.* [76]:

$$\epsilon_r^*(\omega) = \epsilon_r + \frac{\sigma}{j\omega\epsilon_0} + \frac{\chi_1}{1+j\omega\tau} \quad (2.4)$$

where the following parameters are specific to each type of tissue: ϵ_r is the relative permittivity, ϵ_0 is the permittivity of free space, χ_1 is the electric susceptibility, σ is the conductivity and τ is the relaxation time. Finally, $\omega = 2\pi f$ is the angular frequency.

2.3.2. Data-Independent Beamforming

In 1998, Hagness *et al.* [12] implemented a 2D monostatic system based on the Delay And Sum (DAS) beamformer where the breast is illuminated with a UWB pulse and resultant backscattered signals were time-shifted (or delayed) and added (or summed) to produce a synthetic focal point $\mathbf{r} = (x, y)$, which corresponded to each pixel in the image. If a tumour existed at a specific focal point, then the returns from the tumour would add coherently. Returns from clutter due to natural heterogeneous tissue would add incoherently. The energy at this synthetic focus was measured and stored. An energy profile of the breast was created, in which high energy regions indicated the presence of malignant tissue and low energy regions represented normal tissue. If M is the number of monostatic antennas and S_n is the n^{th} backscattered signal, the pixel intensity at \mathbf{r} is given by:

$$I(\mathbf{r}) = \left[\sum_{n=1}^M S_n(\tau_n(\mathbf{r})) \right]^2 \quad (2.5)$$

in which the calculated round-trip time delay between transmit and receive antennas is given by $\tau_n(\mathbf{r}) = 2d_n(\mathbf{r})/vT_s$, where $d_n(\mathbf{r}) = |\mathbf{r} - \mathbf{r}_n|$ is the distance between the n^{th} transmitting antenna at \mathbf{r}_n and the focal point (\mathbf{r}), v is the average velocity of signal propagation in the breast and T_s is the sampling interval between points.

The traditional DAS beamformer was modified by Li *et al.* [70], in 2001, to compensate for attenuation due to radial spreading of the input signal from a transmitting antenna. Li also developed an artifact removal algorithm based on the average of the artifact present in all other channels. In this case, (2.5) becomes:

$$I(\mathbf{r}) = \left[\sum_{n=1}^M w_n B_n(\tau_n(\mathbf{r})) \right]^2 \quad (2.6)$$

where w_n represents the weight used to compensate for radial spreading and B_n represents the backscattered waveform after compensation and artifact removal.

Fear *et al.* [77] developed a similar monostatic system in 3D with patients in the prone position with a cylindrical antenna system as opposed to the planar antenna system described by Hagness *et al.* [12] (and modified by Li *et al.* [70] to accommodate a more realistic breast model). Both cylindrical and planar systems are represented in Figure 2.10. Fear *et al.* [77] included an artifact removal algorithm. In [29], Fear *et al.* tested this approach on both 2D and 3D models. For both their studies, they compensated for the radial spreading of the signals in a cylindrical system using a weighting factor of $1/r$.

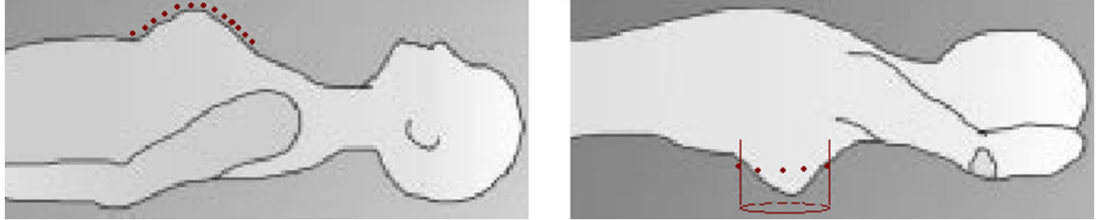


Figure 2.10. The planar antenna configuration is shown on the left when the patient is in a supine position, while the circular antenna configuration is shown on the right when the patient is in a prone position. This image is adapted from [29].

In 2003, Nilavalan *et al.* [72] extended the monostatic DAS beamformer firstly developed by Hagness *et al.* [12] to the multistatic case [72]. In this study each transmitted signal was recorded in all receiving antennas simultaneously, significantly increasing the amount of energy recorded from any scatterer present. The following equation defines the energy return of a volume point of interest, over a window corresponding to the transmit pulse width τ :

$$V = \int_0^\tau \left(\sum_{i=1}^{N(N-1)/2} w_i y_i(t - T_i) \right)^2 dt \quad (2.7)$$

in which N represents the number of multistatic antennas, y_i represents the signals and w_i are factors that are applied to compensate for differences in the predicted attenuation between the round-trip paths (which depend on the depth of a point of interest and the spacing between the considered transmit and the receive antennas in the multistatic scenario). The recorded backscattered signals were spatially focused by time-aligning the signals based on the distance between the transmitting antenna to the point of interest and from this point to the receiving antenna.

Another variant of the simple DAS beamformer is the Delay-Multiply and Sum (DMAS) beamformer developed by Lim *et al.* [78]. This algorithm involved signals being time-shifted, multiplied in pairs and their products summed to form the synthetic focal point. The intensity value of the focal point $I(r_0)$ is obtained by integrating the contributions over a time window W_a :

$$I(r_0) = \int_0^{W_a} S[n] dt \quad (2.8)$$

in which $S[n]$ is the output signal after pairing multiplications, summing and squaring, $W_a = a\Delta t$, and a is an integer representing the number of FDTD time steps. With this algorithm, the reconstructed images provided an improvement in the identification of malignant tumours over the DAS algorithm, and identified tumours as small as 2 mm.

In O'Halloran *et al.* [79], a novel channel-ranked beamformer, based on a simple DAS beamformer, was presented in which different weights were given to the backscattered signals depending on the propagation distance between the antenna locations and the synthetic focal point, so that the effect of dielectrically heterogeneous breasts would be minimised. The authors contended that signals with a longer propagation distance were more likely to encounter heterogeneity and therefore were more prone to incoherent addition, reducing the overall quality of the breast image, therefore extra weighting was attributed to signals with shorter propagation distances while lower weighting was attributed to signals with longer propagation. All signals were processed through a compensation algorithm so that deeper tumours could still be detected with this beamformer. The intensity value of the focal point $I(r_0)$ is obtained by:

$$I(r_0) = \int_0^h \left[\sum_{i=1}^N \sum_{j=1}^N w_{(i,j)} S_{(i,j)}(t - \tau_{(i,j)}) \right]^2 dt \quad (2.9)$$

in which $w_{(i,j)}$ is a channel-dependent weighting factor. O'Halloran *et al.* reported that for a number of different performance metrics, the channel-ranked beamformer outperformed the traditional DAS beamformer allowing for better detection and focusing of small tumours within the breast.

Microwave Imaging via Space-Time (MIST) beamforming was developed by Bond *et al.* [71] and Davis *et al.* [80], using a monostatic system based on a planar antenna configuration. They introduced a new data adaptive algorithm for the artifact removal, where the artifact was estimated in a particular channel by using a filtered combination of artifacts present in all other channels. The residual mean squared

error was then reduced in the region dominated by the early artifact. Overall, the space time beamformer compensated for frequency dependent propagation effects, such as path dependent dispersion and attenuation effects.

Klemm *et al.* [81] presented an Improved Delay And Sum (IDAS) beamformer applied to a 3D FDTD model of the breast. They further developed the traditional DAS beamformer [12] by introducing an additional weight factor QF for each focal point \mathbf{r} within the breast, which allowed for the improvement of image quality. An energy collection curve was calculated by summing radar signals and was normalised by multiplying by the fraction $1/(1 + \sigma_e)$, in which σ_e is the standard deviation of the energy of all radar signals. A second order polynomial ($y = ax^2 + bx + c$) was fitted to the normalised energy collection curve, in which a is the quality factor associated with that point. In particular, it was observed that the performance of the IDAS beamformer was comparable to the performance of a much more complex beamformer such as the data-adaptive Capon beamformer (presented in Sub-section 2.3.3), while it required comparably less computational time.

O'Halloran *et al.* [82] tested the DAS, DMAS and IDAS with three realistic breast FDTD models based on MRIs available on the UWCEM repository [83] with the dielectric properties reported by Lazebnik *et al.* [26-27]. In [82], O'Halloran *et al.* observed that while both IDAS and DMAS significantly outperformed DAS in breast with higher adipose content, the performance of IDAS and DMAS was significantly reduced when applied to dense breasts with higher content of fibroglandular tissue.

In 2010, O'Halloran *et al.* [84] created a system utilising a Quasi-Multistatic MIST beamformer. The monostatic MIST beamformer used in [71, 80] was modified to record signals from all receiving antennas. The signals were firstly coarsely time-aligned, and then passed through a bank of FIR filters which compensated for the frequency-dependent propagation effects. The outputs from the filters were summed and time-gated, and the energy at that particular position was calculated. The focal point of the beamformer was then scanned throughout the breast, and an energy profile of the breast was created. O'Halloran *et al.*'s [84] multistatic MIST approach was tested on a MRI-derived model that included the recently reported dielectric properties of adipose and fibroglandular tissue established by Lazebnik *et al.* [26-27]. Later in [85], the same algorithm was tested for robustness when examining

across a range of potential clinical scenarios: varying location of tumours, different size of tumours, natural variations in dielectric properties and the use of breast models with different proportions of adipose/fibroglandular tissue.

2.3.3. Data-Adaptive Beamforming

Data-adaptive beamforming achieves a maximum reception in a specific direction for each antenna array element using an estimation of the signal arrival (with noise) from a desired direction. Signals from other unwanted directions, such as interference or clutter, are rejected. Unwanted signals are separated from the desired signals by adjusting weights (or steering vectors) at each sensor in the antenna array. In general, data-adaptive beamforming allows for better resolution and better interference suppression capability than data-independent beamformers.

The Standard Capon Beamformer (SCB), developed by Capon *et al.* [86] in 1969, involved adaptively selecting weight vectors to minimise array output power provided that the desired signal did not suffer from distortion effects. The SCB is likely to outperform the data-independent methods, such as those described in Sub-section 2.3.2, as a result of rejecting unwanted signals. However, in reality the steering vector is very sensitive to differences between the assumed and actual signal arrival angles and the performance of the Capon beamformer can be highly compromised.

Several changes have been made to the SCB to improve on the robustness of the algorithm. For example, Li *et al.* [87-88] improved the SCB algorithm by introducing an uncertainty set for the array steering vectors, creating the Robust Capon Beamformer (RCB). The RCB also allowed for a scaling ambiguity when estimating the desired signal power, enhancing previous adaptations such as the Robust Minimum Variance Beamforming by Lorenz *et al.* [89-90], and allowed for simplified and efficient computation of the weight vectors.

Guo *et al.* [91-92] introduced two changes to adapt the RCB for the specific case of breast cancer detection by creating: the Robust Weighted Capon Beamformer (RWCB) and the Amplitude and Phase Estimation (APES) approach. The RWCB used the strategy used in Li *et al.*'s [87-88] RCB algorithm, implementing a weighting strategy to improve performance based on the adjustment of the sample Covariance Matrix. The APES approach specifically assumed the signal waveform is

previously known, estimating energy from the calculated amplitude of the backscattered signal.

Xie *et al.* [93-94] extended the RCB algorithm into 3D and the multistatic approach to create the Multistatic Adaptive Microwave Imaging (MAMI) algorithm. Their MAMI algorithm involved two stages of computation of the RCB algorithm to eliminate errors within the steering vector and making the algorithm more robust. First, the data-adaptive RCB was used spatially to calculate a vector with multiple backscattered waveforms for each probing signal. Second, using the RCB algorithm again, the previously estimated vector of waveforms was used to recover a scalar waveform which was used to compute the backscattered energy.

For the work described in this thesis, the beamformer is based on the DAS principle. This is due to the fact that the DAS algorithm is easy to implement and because the development of an improved beamformer is outside the scope of this thesis in particular. The system used in this thesis is described in detail in Chapter 4.

2.4. Classification Techniques

In addition to using MI to create images of backscattered energy from which the presence of tumours may be inferred, a number of studies have also used more detailed analysis of backscatter in order to classify tumours. In addition to using MI to create images of backscattered energy from which the presence of tumours may be inferred, a number of studies have also used more detailed analysis of backscatter in order to classify tumours. In these studies, it has been confirmed that the morphology of tumours (namely shape, margin, size and density) highly influences the Radar Target Signature of a scattering object detected by UWB radar breast cancer imaging. Furthermore, it is acknowledged that the resonance scattering phenomenon varies with the border profiles of the lesions. Altogether, these findings suggest that benign and malignant tumour classification is feasible. The classification approaches available in the literature are briefly reviewed here for both 2D and 3D systems.

2.4.1. Classification in 2D Systems

Chen *et al.* [95] presented a preliminary study which addresses the effect of the morphology of a tumour mass on microwave signature in 2D UWB radar imaging.

In their study, a 2D breast phantom was modelled with homogeneous normal breast tissue and surrounded by a layer of skin matching the Debye parameters used in [71], and tumour shapes were created based on polygonal approximation of tumour boundaries in X-Ray mammograms by Rangayyan *et al.* [41]. Tumours were modelled with an ellipse baseline which is then modified with two parameters, Q and ΔB , which determine the level of “ruggedness” to recreate whether the tumour is benign or malignant. In this study, a monostatic radar array was used, in which a UWB pulse was transmitted and the backscattered signal was recorded at the same location, one by one for all the antenna elements in the array. As first noted by Huo *et al.* [96-97], the early time response of UWB backscattered signals depends on the specular reflection from the surface of the target and is relatively independent of the mass morphology, while the late-time response is greatly influenced by the Complex Natural Resonances (CNRs) of different mass morphologies and dielectric properties of the targets, and so is directly affected by varying Q and ΔB .

In [98], Chen *et al.* introduced some heterogeneity into the breast phantom by modelling areas of clutter around tumour models, recreating breast phantoms based on the three groups defined by Lazebnik *et al.* [27]; they also created breast phantoms with multiple tumour models. The Fractional-Sequence CLEAN (FS-CLEAN) algorithm was used to examine the late-time response of high-scattering masses within the breast. While they obtained promising results on the detection of malignant tumours in high adipose breast phantoms, the classification of tumours was compromised in breast phantoms with higher content of fibroglandular tissue or when multiple tumours were placed within the breast.

In later studies, Chen *et al.* [45-46] introduced the UWB Multiple-Input Multiple-Output (MIMO), or Multistatic Radar architecture which can transmit multiple probing signals that may be chosen arbitrarily and independently, allowing for improved performance in the detection of suspicious regions when compared to [98]. In this study the tumours were modelled in the same fashion as their previous studies and were grouped into two categories to represent benign and malignant tumours: oval/macrolobulated and microlobulated/spiculated, respectively. Also, all tumour models had the same average size throughout the study. In these studies, robust lesion binary classification was accomplished through two fusion schemes: the Selection-Combining (SC) fusion and Log-Likelihood-Ratio (LLR)-based fusion. On average, for both fusion strategies, the accuracy achieved for shape classification

using a binary classifier was 86.7%.

In [47], Chen *et al.* investigated the feasibility of applying contrast agents to enhance the contrast between tumours and other types of normal tissue within the breast. The agents used in this study were microbubbles and Single-Walled Carbon NanoTubes (SWCNTs), as previously used by Mashal *et al.* [99-100]. Using a monostatic UWB radar, the late-time response of the target, before and after the inclusion of the contrast agents on the previously located tumour sites, was analysed and the differential damping factors of the CNRs were extracted. This study showed that diseased tissues can be correctly classified between 80.0-83.3% for microbubbles, and between 73.3-78.3% for SWCNTs.

In [48], Teo *et al.* looked at the early-time portion of the backscatter breast response for tumour classification, as this provides a higher signal strength. It was also known beforehand that a rough tumour has more significant scattering spicules than a smoother tumour. In their paper, the position of the breast tumour was approximately known beforehand and a set of transceivers (transmit/receive antennas) were located around a portion of the breast. A system with a 2D heterogeneous breast modelled with clutter sources with 30 smooth and 30 spiculated tumours was investigated and the two tumours were well separated for all results.

2.4.2. Classification in 3D Systems

Davis *et al.* [44] addressed the classification of breast tumour models based on their Radar Target Signature (RTS) obtained through Ultra Wideband microwave backscatter. The shape of the tumours was modelled using Gaussian Random Spheres, first introduced by Muinonen [101-102], to represent three different shapes (spiculated, microlobulated and smooth) and four different sizes of tumours (average radius of 2.5, 5, 7.5 and 10 mm).

The different tumour models, as well as a homogeneous breast tissue, were modelled in a 3D Total-Field/Scattered-Field (TF/SF) FDTD model with the dielectric Debye parameters for malignant tissue and for normal homogeneous breast tissue, respectively, as presented by Bond *et al.* [71]. In such a simulation scenario only the response from the tumour is present in the backscattered signals and can be directly analysed. In their paper, Davis *et al.* acknowledged that while the UWB

radar will not provide sufficient resolution to reconstruct details on a millimetre scale, it can potentially provide enough information to classify a tumour by means of its RTS.

In [44], a full classification algorithm was applied to a database of tumour models using the leave-one-out method [103]. The full classification algorithm is composed of two methods for basis selection, Principal Component Analysis (PCA) and Local Discriminant Bases (LDB), and is followed by a Linear Discriminant Analysis (LDA)-based classifier. The system was trained and tested with a 10 dB Signal-To-Noise Ratio. Since the results showed that the size classifier outperformed the shape classifier for the full database, approximately 97 and 70% respectively, a Size-Then-Shape Cascade Classifier was considered as the main classifier for the database of tumour models. A performance of approximately 86% was achieved for a shape classification in the Cascade Classifier.

2.5.Conclusions

This chapter presented background information on tumour cells and how breast cancer is formed, and also on the dielectric properties inherent to breast and tumour tissues. Different beamforming algorithms used in MI based on UWB Radar were reviewed as well as methods for classification of different types of tumours. The material presented in this chapter motivates the simulation experiments developed in the following chapters:

- in Chapter 3 the modelling of tumours of different sizes and shapes is addressed;
- in Chapter 4 the comparison between planar and circular antenna configurations, as well the optimisation of the number of antennas in both systems, is presented;
- in Chapters 5 and 6, different classification algorithms are developed for classification of large tumour databases embedded in homogeneous and heterogeneous breast models, respectively.

The next chapter discusses the methodology used for tumour modelling for the work described in this thesis.

Chapter 3

Tumour Modelling

3.1.Introduction

This chapter discusses the methodology used to model benign and malignant tumours for the work subsequently described in this thesis. To aid in the development of classification methodologies, it is necessary to use a database of accurate numerical tumour models which incorporates different sizes, growth patterns from benign to malignant and corresponding dielectric properties as well as classification algorithms that further detect and differentiate between benign and malignant tumours.

Gaussian Random Spheres (GRSs) are used to model benign and malignant tumours, following Muinonen's [101] algorithm originally designed for modelling asteroids and comets in an astrophysics context. The reason the GRSs are used to model tumour growth and development is that they can be easily modified to provide different sizes, shapes and textures of surface in 3D, as these are characteristics that most significantly influence the Radar Target Signature (RTS) of tumours which are used for classification purposes. As noted in the previous chapter, benign tumours typically have smooth surfaces and have spherical, oval or at least well-circumscribed contours. Conversely, malignant tumours usually present rough and complex surfaces with spicules or microlobules, and their shapes are typically irregular, ill-defined and asymmetric. Although size may play an important role when analysing the development of a tumour over a period of time, and may be an indication of malignancy, in this study the primary concern is the analysis of small tumours (up to 1cm in radius), therefore, shape and texture of the surface of a tumour are the two most important characteristics that will help differentiate between a benign and a malignant tumour [41-43, 57].

3.2. Basis of the Gaussian Random Sphere Method

The mathematical model for GRSs was developed and presented by Muinonen [101-102, 104], and later adjusted to the breast tumour context by Davis *et al.* [44]. The shape of the GRS is given by a radius vector, $\mathbf{r} = \mathbf{r}(\vartheta, \varphi)$, which is defined by the logradius, or logarithmic radius, $s = s(\vartheta, \varphi)$:

$$\mathbf{r}(\vartheta, \varphi) = \alpha \exp \left[s(\vartheta, \varphi) - \frac{1}{2} \beta^2 \right] \quad (3.1)$$

$$s(\vartheta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^l s_{lm} Y_{lm}(\vartheta, \varphi) \quad (3.2)$$

In the equations above (ϑ, φ) stand for the spherical coordinates, α is the mean radius, β is the standard deviation of the logradius, Y_{lm} are the orthonormal spherical harmonics, s_{lm} are the spherical harmonics weight coefficients, in which l and m stand for the degree and order of the expansion.

Furthermore, the covariance functions of the radius and the logradius, $\alpha^2 \Sigma_r$ and Σ_s , respectively, and the corresponding variances, $\alpha^2 \sigma^2$ and β^2 , are interrelated through:

$$\Sigma_r = \exp(\Sigma_s) - 1 \quad (3.3)$$

$$\sigma^2 = \exp(\beta^2) - 1 \quad (3.4)$$

in which σ represents the standard deviation of the radius. The covariance function of the logradius can be further given by:

$$\Sigma_s = \beta^2 C_s(\gamma) \quad (3.5)$$

in which, γ is the angular distance between two directions (ϑ_1, φ_1) and (ϑ_2, φ_2) , and C_s is the logradius correlation function. After adjusting for the degree of the expansion (l) as in [102], the correlation length ℓ and the correlation angle Γ are defined by:

$$\ell = \frac{1}{\sqrt{-C_s^{(2)}(0)}} \quad (3.6)$$

$$\Gamma = 2 \arcsin \left(\frac{1}{2} \ell \right) \quad (3.7)$$

in which $C_s^{(2)}$ is the logradius correlation function in which $\Sigma_s^{(2)}$ is the second derivative of the covariance function with respect to $\gamma = 0$. GRSs can be modified mathematically to model both malignant and benign tumours of different sizes by varying the mean radius (α) and the covariance function of the logradius (Σ_s) [102].

The basis for the GRS method implemented in this work was taken from the ‘‘G-sphere’’ software developed by Muinonen [104], translated into the Matlab

programming language.

Four different models of tumours at four different sizes are considered in this chapter. Malignant tumours are represented by spiculated and microlobulated GRSs. In contrast with Davis *et al.* [44], benign tumours are divided into two type categories: smooth and macrolobulated. The aim of having two categories for benign tumours is twofold: it becomes possible to differentiate the stage of development of a benign tumour and a stage of pre-malignancy may be detected sooner, and, additionally, the number of sub-categories for malignant shapes and for benign shapes is equal, allowing for better training of classifiers.

Microlobulated, macrolobulated and smooth GRSs are obtained by varying the correlation angle (Γ) discretely in intervals of 5° . The correlation angle varies approximately between 5° and 20° for microlobulated GRSs, between 25° and 45° for macrolobulated GRSs and between 50° and 90° for smooth GRSs. Smooth GRSs have a spherical or ellipsoid appearance with light distortions on the surface. Macrolobulated GRSs often resemble ellipsoid figures in which one of the dimensions is significantly higher than the other or show large lobular protuberances. Microlobulated GRSs are often represented by spherical volumes with several smaller lobular protuberances of different sizes on the surface.

Spiculated GRSs are obtained by adding 3, 5 or 10 spicules to smooth GRSs, as previously used in [44]. The increasing number of spicules reflects the evolution of the tumour model as it suggests that the tumour is spreading in more directions. The spicules are modelled with cones of height 2 cm and with a radius matching the average radius of the smooth GRSs to which they are added. Consequently, this implies that the spicules are more prominent for GRSs with smaller average radius than for GRSs with larger average radius, as the length of the spicules remains constant. Finally, the centre of the GRSs matches the centre of the base of the spicules which are randomly directed to any direction.

The average radius of all types of spheres are 2.5, 5, 7.5 or 10 mm, as previously used for the database in [44].

3.3.Tumour Models

For the work described in this thesis, a database of 480 tumour models is created

including 30 samples for each of the four sizes and the four shapes. Examples of benign tumour models based on the GRSs method, with radii of 5 and 10 mm, are shown in Figure 3.1 and Figure 3.2.

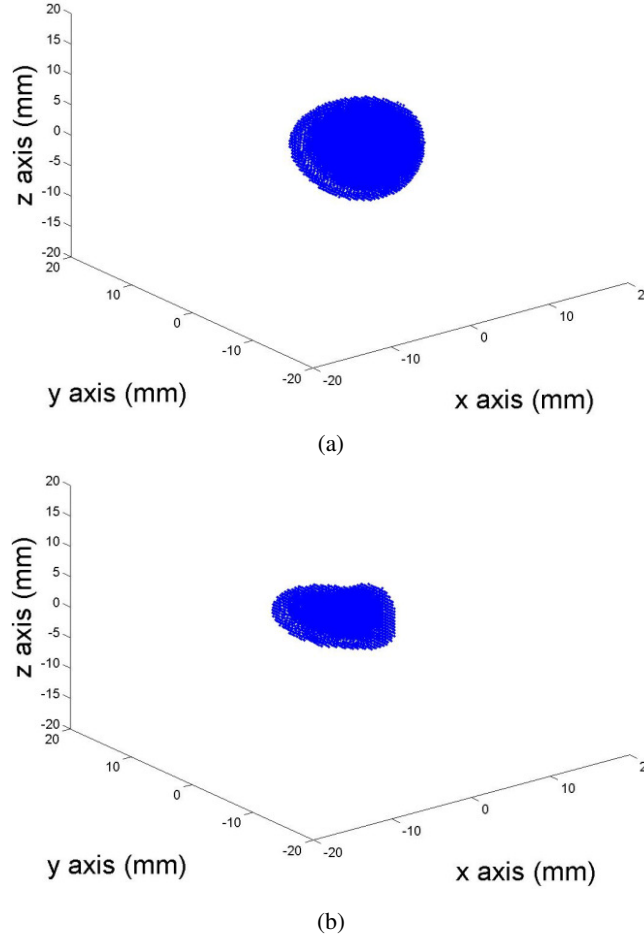
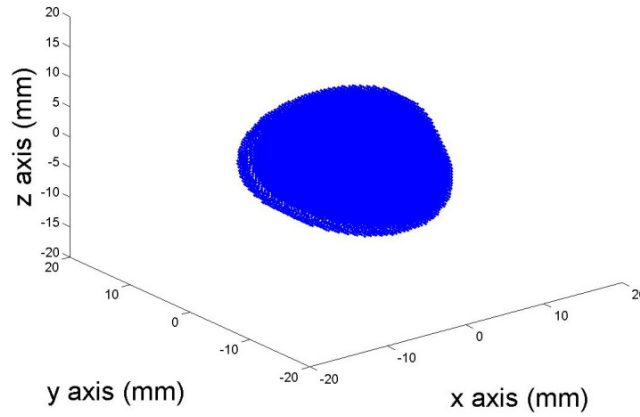
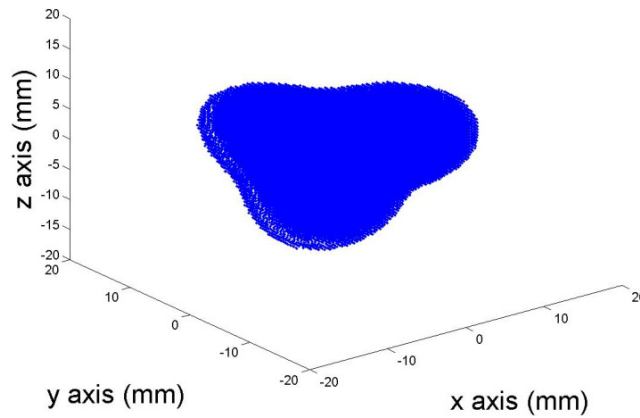


Figure 3.1. Samples for different Gaussian Random Spheres (GRSs) representing benign tumour models. From top to bottom: (a) smooth and (b) macrolobulated. The GRS models have an average radius size of 5 mm.



(a)



(b)

Figure 3.2. Samples for different Gaussian Random Spheres (GRSs) representing benign tumour models. From top to bottom: (a) smooth and (b) macrolobulated. The GRS models have an average radius size of 10 mm.

Examples of malignant tumour models, created with the GRS method, with radii of 5 and 10 mm are shown in Figure 3.3 and Figure 3.4. In particular, a microlobulated GRS is presented as well as spiculated GRSs with: 3, 5 and 10 spicules.

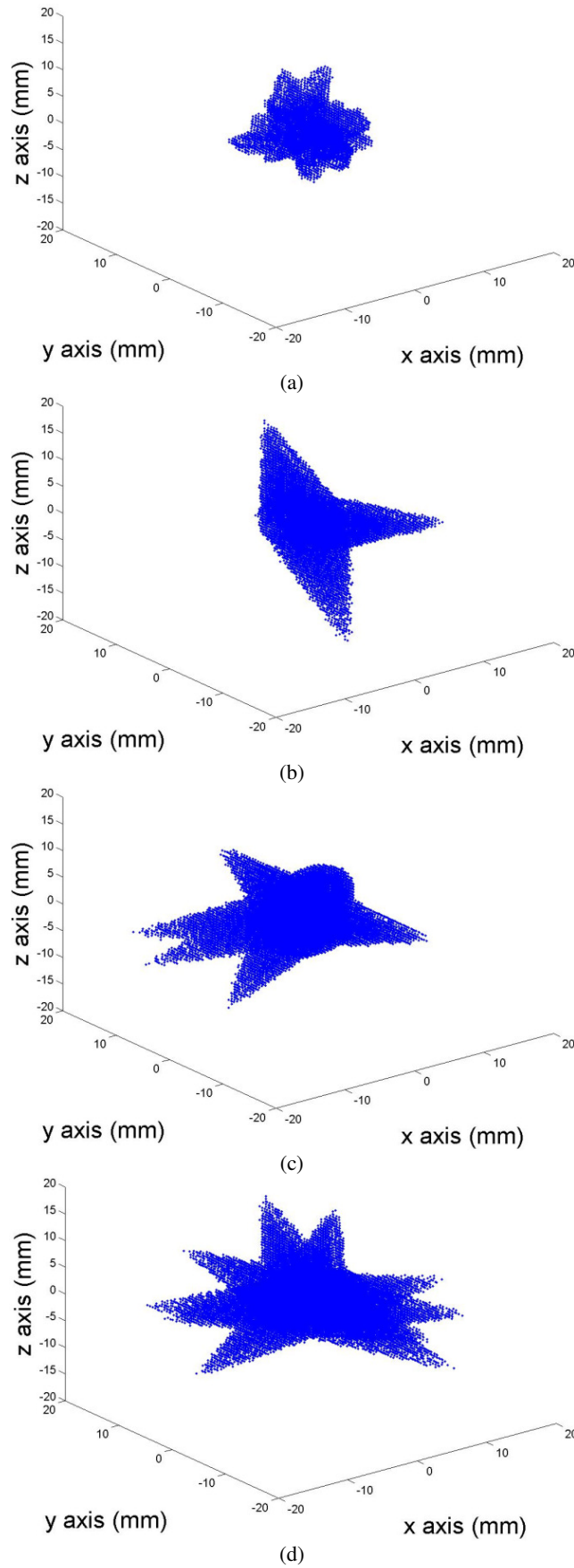


Figure 3.3. Samples for different Gaussian Random Spheres (GRSs) representing malignant tumour models. From top to bottom: (a) microlobulated, (b) spiculated with 3 spicules, (c) 5 spicules and (d) 10 spicules. The GRS models have an average radius size of 5 mm.

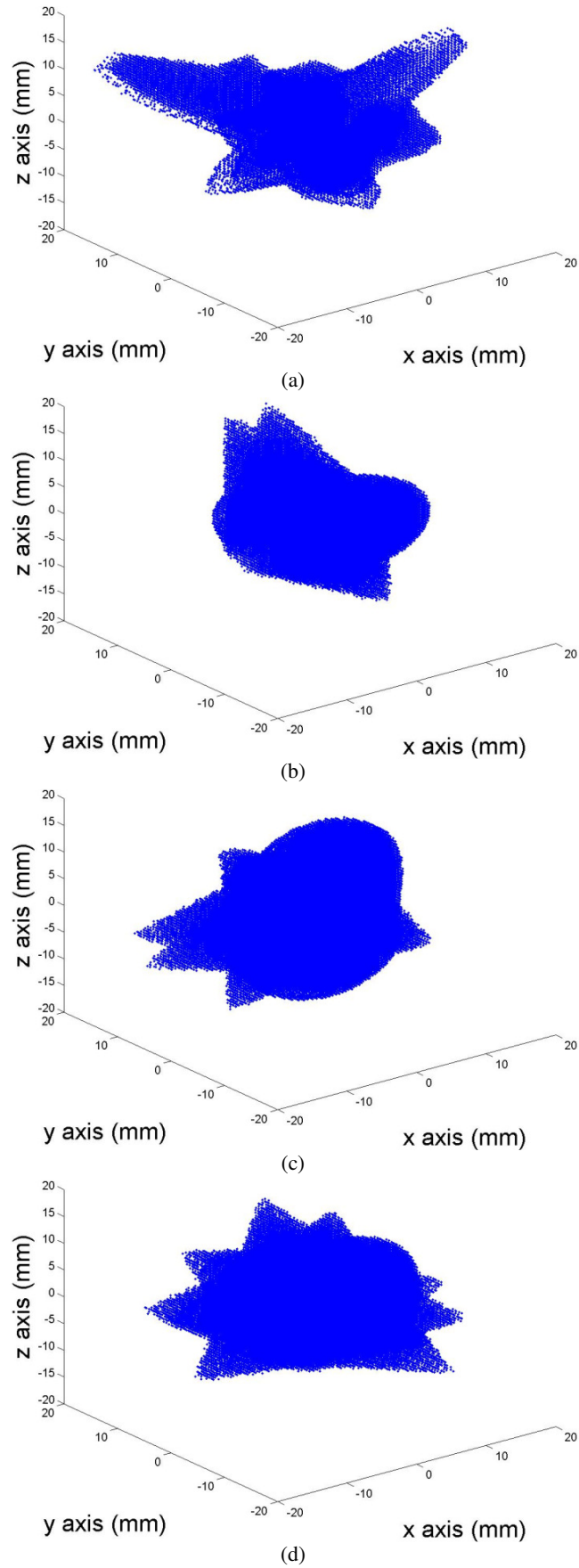


Figure 3.4. Samples for different Gaussian Random Spheres (GRSs) representing malignant tumour models. From top to bottom: (a) microlobulated, (b) spiculated with 3 spicules, (c) 5 spicules and (d) 10 spicules. The GRS models have an average radius size of 10 mm.

All surface and interior points of the GRS models are then modelled with the Debye parameters using the single-pole formulation, as mentioned in (2.4) , repeated here for completeness:

$$\epsilon_{\infty}^*(\omega) = \epsilon_{\infty} + \frac{\sigma_s}{j\omega\epsilon_0} + \frac{\Delta\epsilon}{1 + j\omega\tau}$$

The parameters for malignant tissue, as established by Lazebnik *et al.* [27] are the following: $\epsilon_{\infty} = 6.749$, $\Delta\epsilon = 50.09$, $\sigma_s = 0.794 \text{ Sm}^{-1}$ and $\tau = 10.50 \text{ ps}$.

The numerical simulations detailed in Chapter 5 allow for the acquisition of the backscattered response of tumours embedded in a homogeneous breast model. As an example, the backscattered signals of a smooth tumour with different values of radius, and the backscattered signals of the smallest tumours (with 2.5 mm) with different shapes are presented in Figure 3.5 to illustrate the differences between the signals produced by tumours of different sizes and shapes.

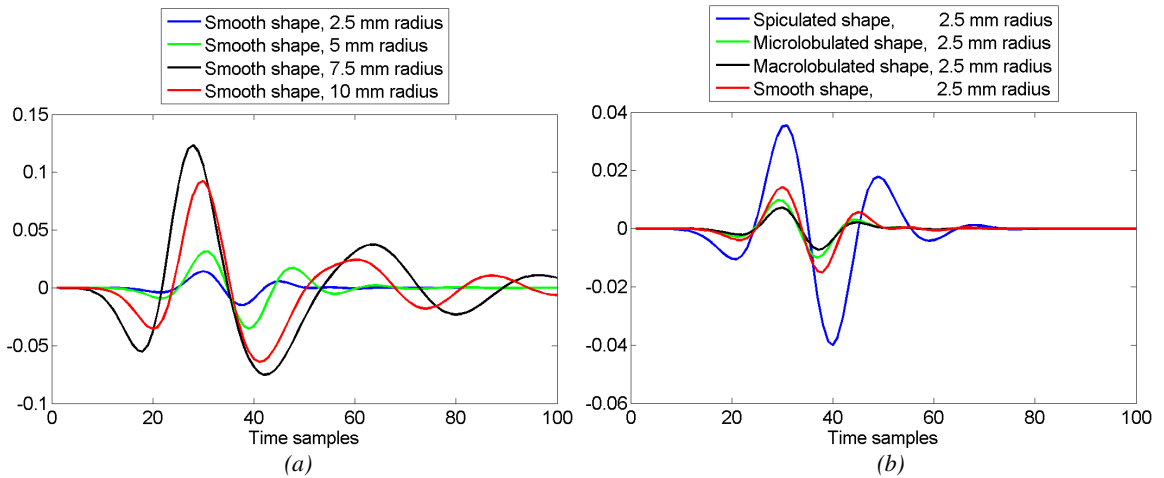


Figure 3.5. Sample of (a) backscattered signals for tumours of identical shape with different sizes and (b) backscattered signals for tumours of different shapes with identical size.

3.4. Conclusions

This chapter has discussed the modelling of tumours, which will be needed for simulation purposes in the following chapters, through the GRSs method. A database of 480 realistic tumour models based on GRSs was created, including four different shapes (two benign and two malignant) and four different sizes. With this database it was possible to recreate various stages of growth and development of tumours from

2.5 to 10 mm in radius size and between different degrees of benignity and malignancy. These variations are expected to influence the Radar Target Signature of tumours significantly enough to allow for the different sizes and shapes to be successfully classified.

The next chapter will address the optimisation of the UWB system by comparing two different antenna configurations: the planar and the circular systems. The two antenna configurations are compared and evaluated using a number of performance metrics, and in a number of different conditions.

Chapter 4

Comparison of Planar and Circular Antenna Configurations

4.1.Introduction

This chapter is motivated by the need to optimise the UWB system design for breast cancer imaging, in particular the antenna configuration. Different antenna configurations have been used in different studies, particularly planar and circular configurations.

In this study, planar and circular antenna configurations are compared in a range of conditions. Firstly, both configurations are evaluated with tumours varying in size from 2 to 10 mm in diameter. In addition, the antenna configurations are evaluated for tumours placed in different locations of the breast. In each case, the optimum number of antennas for both antenna configurations is investigated. The robustness of the antenna configurations is investigated by increasing the dielectric variance in normal breast tissue. Evaluation is carried out using a range of different performance metrics that measure the ability of the imaging system to highlight tumours in the presence of clutter, and also evaluate the localisation ability of the system.

The findings hereby presented have been published in the following papers:

- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Comparison of Planar and Circular Antenna Configurations For Breast Cancer Detection Using Microwave Imaging", *Progress In Electromagnetics Research*, Vol. 99, pp. 1-20, 2009 [49].
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Antenna configurations for Ultra Wide Band Radar Detection of Breast Cancer", in *Proceedings of the SPIE BIOS West*, San José, CA, USA, Vol. 7169, No. 9, pp. [71691M, 12], 2009 [58].

4.2.Existing Antenna Configurations

Different antenna configurations have been used in a number of studies regarding UWB Radar Microwave Imaging. In this study two main system configurations are considered: the planar configuration initially developed by Hagness *et al.* [12] and the circular configuration developed by Fear *et al.* [29].

Each antenna configuration is defined by the orientation of the patient and the position of the antenna array. In the planar configuration, the patient is oriented in the supine position with a planar antenna array placed across the naturally flattened breast. This configuration has been used by Hagness *et al.* [12], Bond *et al.* [71], Davis *et al.* [80], Nilavalan *et al.* [72], Li *et al.* [70] and O'Halloran *et al.* [84].

Conversely, in the circular configuration, the patient lies in the prone position with the breast naturally extending through an opening in the examination table, while a circular array of antennas surrounds the breast. The circular configuration has been used by Fear *et al.* [77, 105-106], Xie *et al.* [93-94], Klemm *et al.* [73, 81] and Craddock *et al.* [107]. In the literature, this configuration is often referred to as cylindrical or hemispherical due to its 3D configuration. In a cylindrical configuration, a circular array of antennas is placed at a certain depth of the breast to analyse the corresponding cross section through the breast; the same array of antennas is then vertically translated at different depths to further analyse the breast [77, 105]. For the hemispherical configuration, a circular array of antennas is placed around the base of the breast and then vertically translated at other depths towards the nipple, however for this configuration the antennas are kept at a constant distance from the skin [73, 81, 93-94, 106-107].

Both planar and circular antenna configurations are shown in Figure 2.10, and represented again below for completeness, for patients in supine and prone positions, respectively.



Figure 2.10. The planar antenna configuration is shown on the left when the patient is in a supine position, while the circular antenna configuration is shown on the right when the patient is in a prone position. This image is adapted from [29].

4.3.Simulation Procedure and Performance Metrics

This sub-section describes the simulation procedure used for numerical modelling of the breast, as well as the performance metrics used to compare the performance of the planar and circular antenna configurations.

4.3.1.Numerical Model of the Breast

In order to test both antenna configurations, a numerical model of the breast is created for both configurations. A 2D FDTD model of the breast, similar to the model used by Hagness *et al.* [12] is used to examine the planar configuration, while a separate 2D breast model, as used by Fear *et al.* [77], is used to test the circular configuration. In both models, a 2 mm layer of skin covers the breast.

As noted in Chapter 2, Debye parameters are used to model the frequency-dependent propagation characteristics of the various tissues. The Debye parameters for skin are chosen to fit published data by Gabriel *et al.* [23], while the Debye parameters for malignant tissue are those used by Bond *et al.* [71]. Normal breast tissue is modelled with the Debye parameters for adipose tissue [12, 70, 77]. To account for the dielectric differences between fibroglandular and adipose tissues, a dielectric variation is randomly incorporated in a checkerboard fashion, using a Gaussian distribution, to 4-mm-side squares within the 2D model of the breast, as used in [12, 77]. The checkerboard pattern is randomly created for all simulations. Models of the breast are created to incorporate the following dielectric variations, which represent the variation within the normal breast tissue between adipose and fibroglandular tissue: $\pm 10\%$, $\pm 20\%$, $\pm 30\%$, $\pm 40\%$ and $\pm 50\%$, where these variations are with respect to the nominal dielectric properties calculated according to the Debye parameters for adipose tissue. The Debye parameters for each type of tissue, along with the relative permittivity and conductivity at the centre frequency, are shown in Table 4.1.

Table 4.1. Debye parameters for the FDTD model and dielectric properties of each tissue at the centre frequency of the input pulse.

Tissue	ϵ_∞	$\Delta\epsilon$	σ_s (sm^{-1})	τ (ps)	Relative Permittivity	Conductivity (sm^{-1})
Skin	15.63	8.2	0.82	12.6	21.65	2.35
Tumour	7	47	0.15	7	49.2	6.1
Adipose	3.20	1.65	0.035	16	4.30	0.38

For the supine (planar) examination, the breast is naturally flattened and the antennas lie directly on the skin across a span of 80 mm. As the breast flattens, it is assumed that the maximum depth of the breast tissue is 42 mm [12]. The antenna array is backed by a synthetic material matching the dielectric properties of skin. The antenna array elements are modelled as electric-current sources. For the prone (circular) examination, the antennas form a circle around the breast. The radius of the breast is 36 mm, including the 2 mm layer of skin [77]. The breast and antenna array are once again backed by a synthetic material matching the dielectric properties of skin. A representation of both models is shown in Figure 4.1; these diagrams also show the co-ordinate system used for locating tumours in the breast.

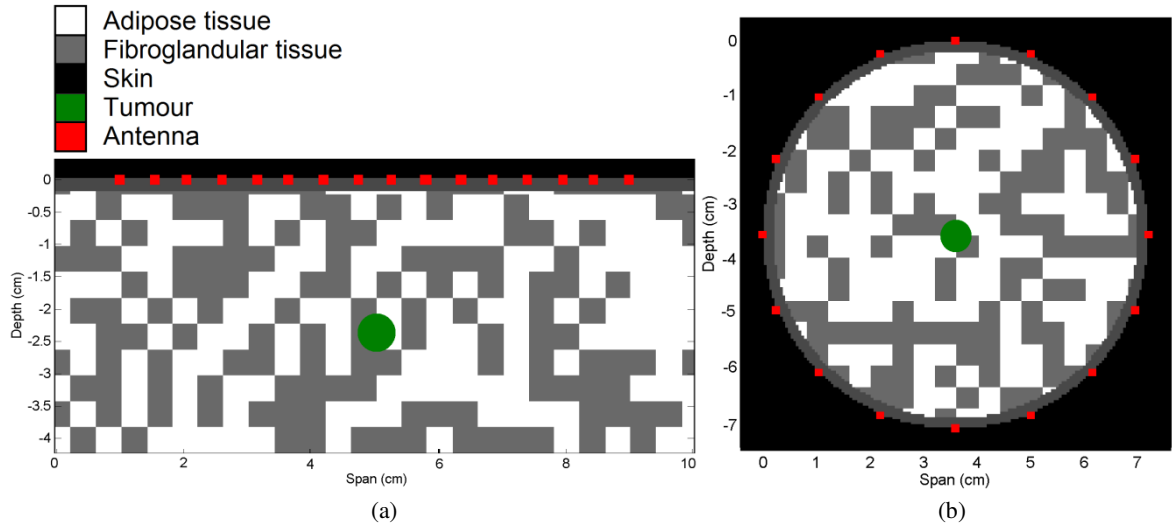


Figure 4.1. The supine and prone breast models are shown in (a) and (b), respectively. In the grid area, the white regions represent adipose tissue, while the darker regions represent fibroglandular tissue. The tumour is represented in green. A 2 mm layer of skin limits both models (shown in dark grey), and on top of it the antenna locations are shown in red dots.

The FDTD grid resolution, dx , is 0.5 mm and the time step, dt , is calculated as 0.833 ps (by means of (2.3)). A scan involves sequentially illuminating the breast model with a UWB pulse from each antenna, while recording the backscattered signal at the same antenna. For the tests in this chapter the number of antennas varies

between 8 and 20. Before further processing, the signals are downsampled from 1200 GHz (the sampling frequency of the FDTD simulation) to 50 GHz. The input signal is a 150 ps differentiated Gaussian pulse, with a centre frequency of 7.5 GHz and a -3dB bandwidth of 9 GHz. The time-domain and magnitude spectra of the UWB input pulse are represented in Figure 4.2.

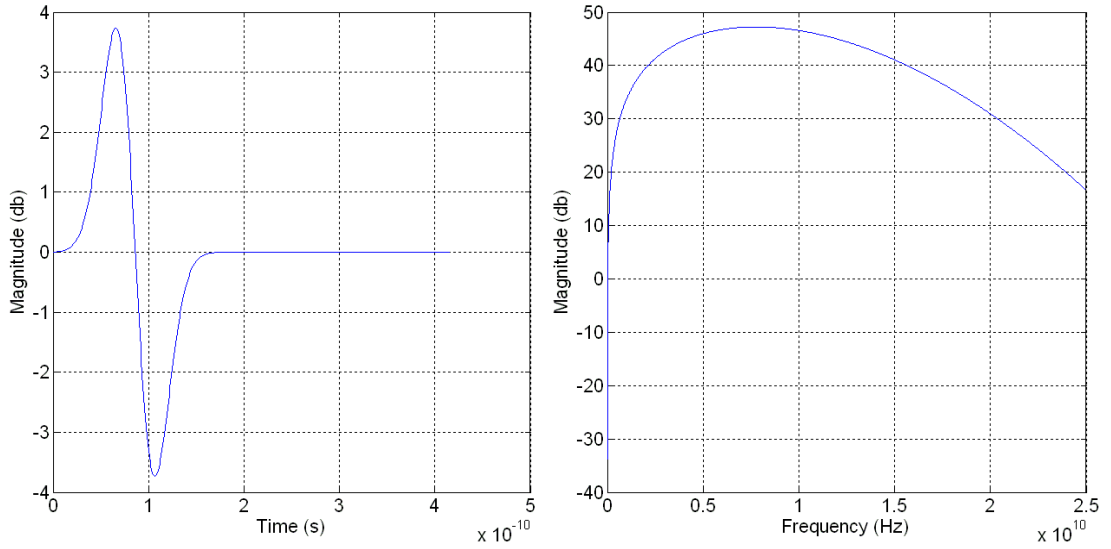


Figure 4.2. The time-domain (left) and magnitude (right) spectra of the UWB input pulse.

4.3.2. UWB Imaging System

A skin-breast artifact removal algorithm is applied to the recorded backscattered signals to remove the input signal and the reflection due to the skin-breast interface. Two different skin-breast artifact removal algorithms are used for the two parts of the experimental setup, as will be detailed later in this chapter. The first skin-breast artifact removal algorithm used is based on the algorithm described in Bond *et al.* [71]. Bond's artifact removal is applied, channel-by-channel, by estimating the artifact in each channel as a filtered combination of the artifacts in all the signals from the remaining channels. A simplified block diagram of Bond's artifact removal is presented in Figure 4.3, in which $b_i[n]$ represents discretised backscattered signal recorded in antenna i , and $x_i[n]$ represents the artifact-free backscattered signal.

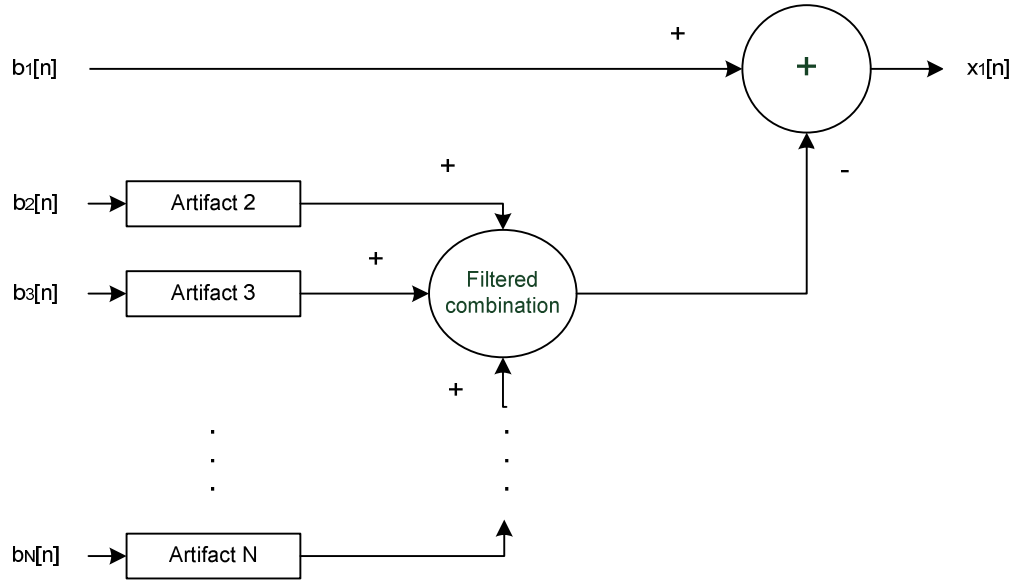


Figure 4.3. Simplified block diagram representing the artifact removal algorithm as described in Bond et al. [71].

The second artifact removal algorithm used is an “ideal” artifact removal algorithm. The “ideal” artifact removal is applied by subtracting the backscattered signals from a homogeneous FDTD model without any tumour from the backscattered signals obtained from the heterogeneous FDTD model containing a tumour, channel by channel, therefore maintaining the effects from the breast heterogeneity and tumour in the backscattered signals, while removing the input signal and the skin-breast artifact.

Finally, a monostatic Delay And Sum (DAS) beamformer is used to create the image of the breast. For the simple DAS beamformer used here, backscattered signals recorded by the antennas are time-shifted, scaled by a scaling factor k (which allows for the tumour response to have similar magnitudes independent of the antennas by accounting for the distance between a possible tumour and the antennas), summed and windowed (with a 5 mm window) to produce a synthetic focal point. The returns from high scattering regions in the breast, such as tumours, will add coherently, whereas returns from clutter, which result from the existence of different normal breast tissues within the breast, will add incoherently. In Figure 4.4, a block diagram of the simple delay and sum beamformer is presented.

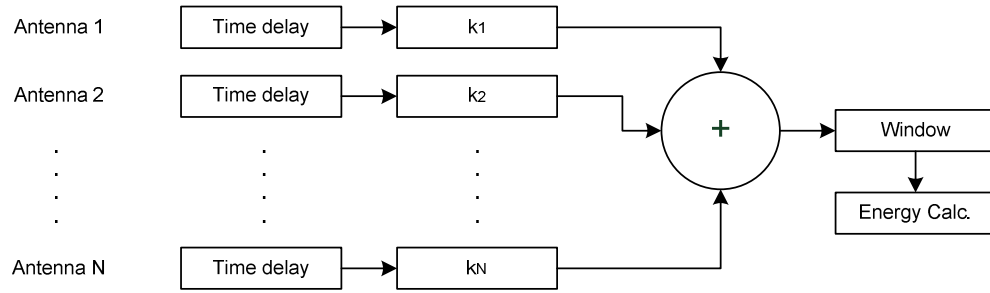


Figure 4.4. Block diagram illustrating the simple delay and sum beamformer for a system with N antennas. The signals recorded in each antenna are time delayed, scaled by a scaling factor k and then added together, finally the signals are windowed and the energy profile is calculated.

4.3.3. Performance Metrics

In order to evaluate the performance and robustness of each antenna configuration, given different tumour sizes and locations within the breast, a number of different metrics are used:

- Signal-to-Clutter Ratio within-breast (SCRwb) [49, 58, 106, 108];
- Signal-to-Clutter Ratio between-breast (SCRbb) [29, 58, 84, 106, 108];
- Signal-to-Mean Ratio (SMR) [49, 78];
- The difference between the actual location of the tumour and the location of the energy peak in the resulting image of backscattered energy (M_{diff}) [49, 58, 84];
- The Full Width Half Maximum (FWHM), which estimates the physical extent of the tumour response [49, 58, 106, 108].

The SCRwb compares the maximum tumour response to the maximum clutter response in the same image of backscattered energy. To obtain the value of the maximum clutter, the maximum pixel value of the image is found in the area of the breast that excludes the tumour peak response up to twice the extent of the FWHM value [49, 58, 106, 108].

The SCRbb compares the maximum tumour response with the intensity at the same location in an image formed with the modelled breast with no tumour. In a real life examination, the potential image without a tumour could be acquired from the opposite breast, as in principle both breasts from a patient should be similar in terms of average dielectric properties, though with different tissue configurations, and it is unlikely that both breasts will have tumours in the exact same location [29, 58, 84, 106, 108].

The SMR compares the maximum tumour response with the mean response of the different tissues across the breast in the same image of backscattered energy [49, 78].

The FWHM measures the distance between the energy peak response of the detected tumour and the point at which the energy of the peak response drops to half [49, 58, 106, 108]. This metric, as well as M_{diff} , determines the ability of the beamformer to effectively localise the tumour within the breast.

4.4.Simulation Results and Discussion

The experiment is developed in two stages which are presented in the following two Sub-sections, 4.4.1 and 4.4.2. The optimisation of the number of antennas is carried out in two situations, when varying the size of tumours, and when varying the location of tumours. The two antenna configurations are then tested for robustness by increasing the variance within the normal tissues of the breast.

4.4.1.Optimum Number of Antennas for Varying Tumour Size

The first experiment involved the optimisation of the number of transmitting/receiving antennas that allow for improved performance. The planar and the circular configurations were tested with a tumour located at the following (span, depth) locations: (5,-2.4) cm and (3.6,-3.6) cm, respectively. The size of the tumour varied from 2 to 10 mm, in steps of 2 mm. Backscattered signals were processed with Bond's artifact removal algorithm and a DAS beamformer. The performance for each antenna configuration was analysed and compared for a varying number of transmitting/receiving antennas (8 to 20), in a total of 650 FDTD simulations with different levels of dielectric heterogeneity, by visual inspection of backscattered images and by using performance metrics.

A sub-set of the metric results, SCRwb and SCRbb, is presented in Figure 4.5, when 6 mm tumours are embedded within the normal breast tissue, with a $\pm 10\%$ dielectric variation.

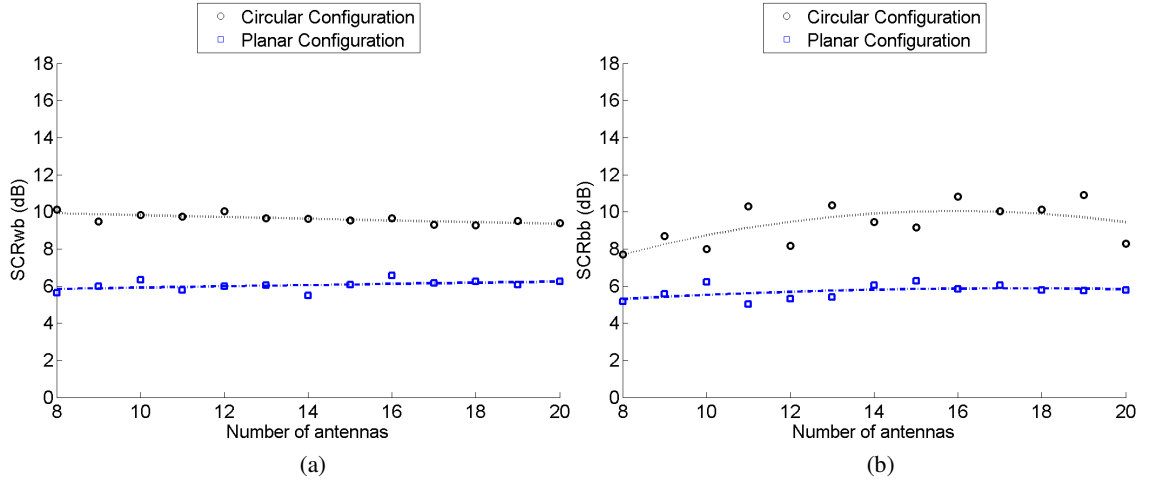


Figure 4.5. Absolute values for the performance metrics: (a) SCR_{wb} and (b) SCR_{bb} versus the number of antennas, with the presence of a 6 mm tumour with a dielectric variation of the normal breast tissue by $\pm 10\%$. Quadratic curves were fitted to the data. For SCR_{wb} the quadratic equations are given by $y = 0.0003x^2 - 0.0557x + 10.3477$ (circular) and $y = -0.0004x^2 + 0.0437x + 5.5195$ (planar). For SCR_{bb} the quadratic equations are given by $y = -0.0370x^2 + 1.1832x + 0.5994$ (circular) and $y = -0.0068x^2 + 0.2333x + 3.8701$ (planar).

From Figure 4.5 and all images of backscattered energy (data not represented), for both antenna configurations, improvement is noticed when 16 to 20 antennas are used. Therefore the optimum number of antennas used is 16 antennas for both antenna configurations, as this number allows for successful detection of tumours without additional computational load. Increasing the number of antennas beyond this point will not increase spatial diversity significantly and therefore will not affect performance significantly. 16 antennas was also found to be the optimum number for all tumour sizes; however for brevity only results for the 6 mm are plotted here, by way of example.

The resulting backscattered images using 16 antennas with tumours from 2 to 10 mm are shown in Figure 4.6, for the circular antenna configuration, and in Figure 4.7, for the planar antenna configuration.

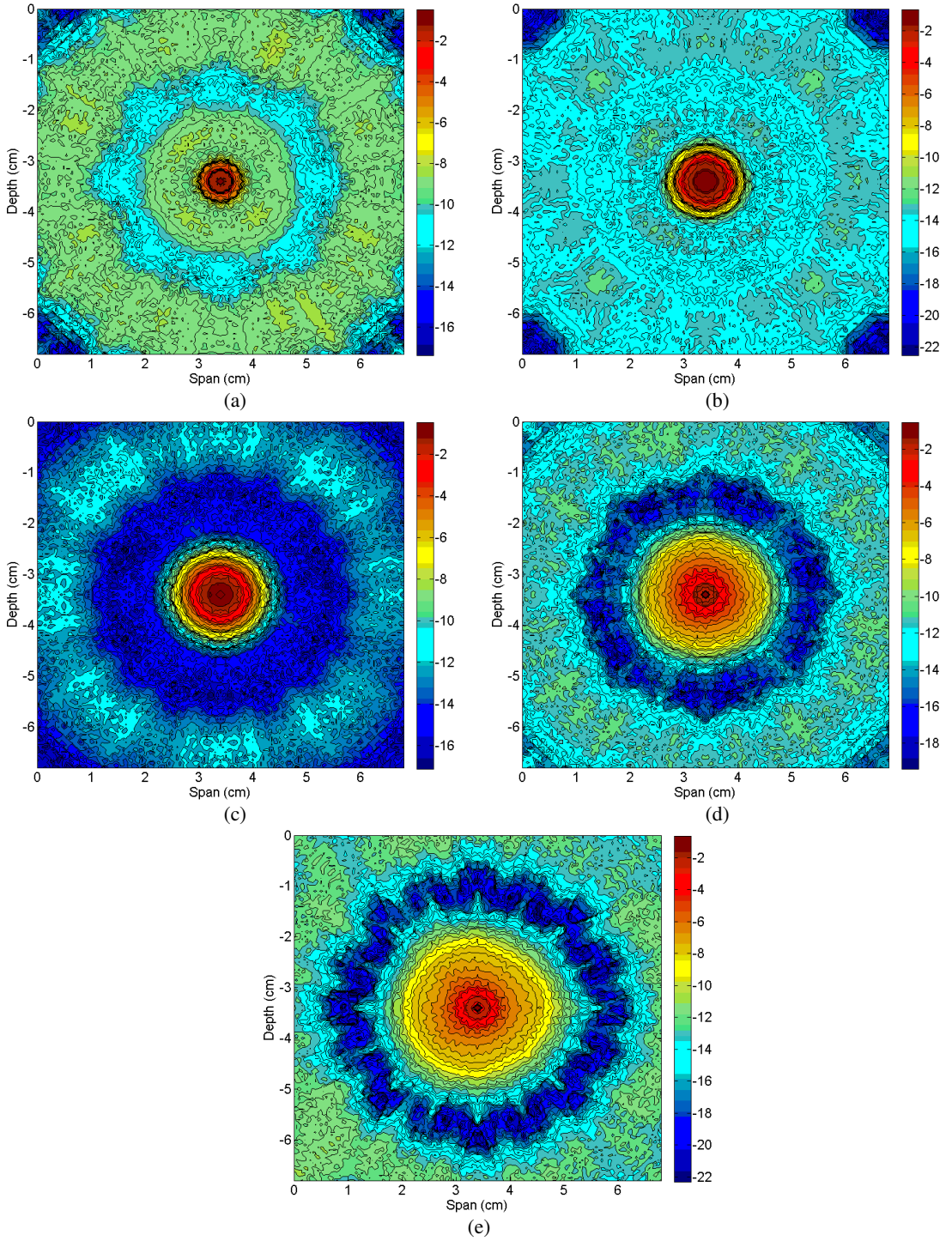


Figure 4.6. Images of backscattered energy (on a dB scale) for the circular antenna configuration, using 16 antennas. For all images, there is a tumour present with its centre at (3.6,-3.6) cm. The size of the tumours is the following: (a) 2mm, (b) 4mm, (c) 6mm, (d) 8mm, (e) 10mm. Note different intensity scales for the different plots.

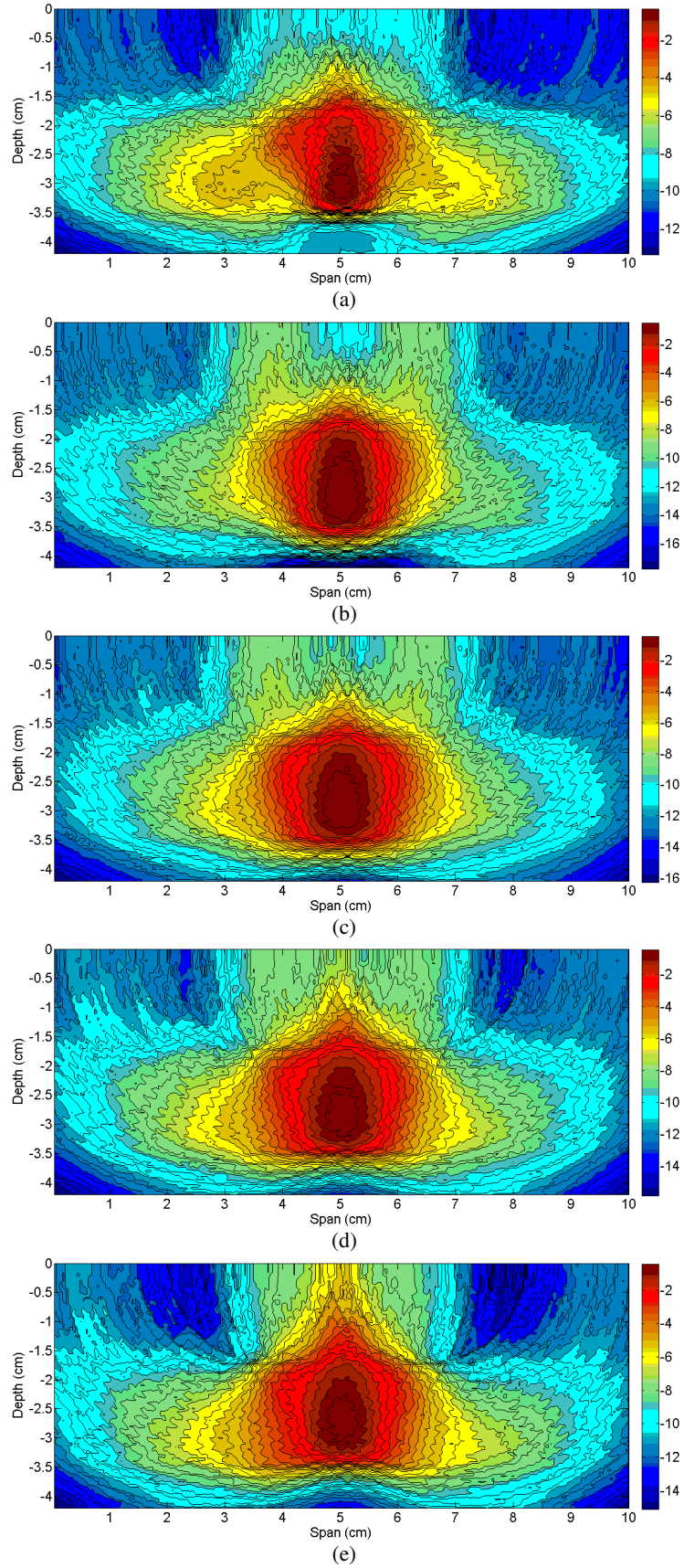


Figure 4.7. Images of backscattered energy (on a dB scale) for the planar antenna configuration, using 16 antennas. For all images, there is a tumour present with its centre at (5,-2.4) cm. The size of the tumours is the following: (a) 2mm, (b) 4 mm, (c) 6 mm, (d) 8 mm, (e) 10 mm. Note different intensity scales for the different plots.

For all considered tumours, it is observed that their positions are easily detected in Figure 4.6 and Figure 4.7. Comparing the two configurations, the circular configuration gives a clearer and more isolated position of the tumour when compared to the equivalent results for the planar configuration. Also in the circular configuration, the centre of the tumour is closer to its original position when compared to the planar configuration for which the tumour seems to be closer to the skin layer than it actually is, suggesting that having antennas around the tumour allows a better focusing of its location when compared to having antennas on only one side of the tumour.

The performance metrics for both antenna configurations using 16 antennas is shown in Table 4.2.

Table 4.2. Summary table with performance metrics results for both antenna configurations with different tumour sizes when using 16 transmitting/receiving antennas.

Metrics	Antenna Configuration	Tumour Size (mm)				
		2	4	6	8	10
SCRwb	Circular	5.8	10.7	9.6	8.4	7.2
	Planar	4.0	7.3	6.6	5.9	5.3
SCRbb	Circular	7.2	11.4	10.8	5.1	4.4
	Planar	4.3	7.3	5.8	5.4	4.5
M_{diff}	Circular	0.5	0.5	0.5	0.5	0.5
	Planar	2.1	3.5	3.5	6.0	6.0
FWHM	Circular	2.5	4.3	6.0	3.9	1.5
	Planar	8.4	9.4	10.1	10.4	11.1

The contrast between the tumour and the breast is higher for the circular than the planar configurations, as indicated by the SCRwb and SCRbb. An exception is for 8 and 10 mm tumours for the SCRbb metric, although the results still remain comparable to those for the planar configuration. The tumour is better localised and “defined” for the circular than the planar configurations, as indicated by metrics M_{diff} and FWHM (M_{diff} is the metric that better reflects the contrast between both antenna configurations, which is supported by visual inspection).

It should be noted that the values for SCRwb and SCRbb first increase when the tumour models increases from 2 to 6 mm, showing an improvement on the performance of the imaging configurations as the tumour shows a stronger response comparatively to clutter. However, when the tumour has a larger dimension of 8 and then 10 mm, these metrics drop in value. This is likely due to the fact that the

window used to isolate the tumour response is optimised for smaller tumours (the window size is 5 mm) and when larger tumours are detected, the window does not cover the entire tumour and, consequently, the system may consider that the rest of the tumour is clutter.

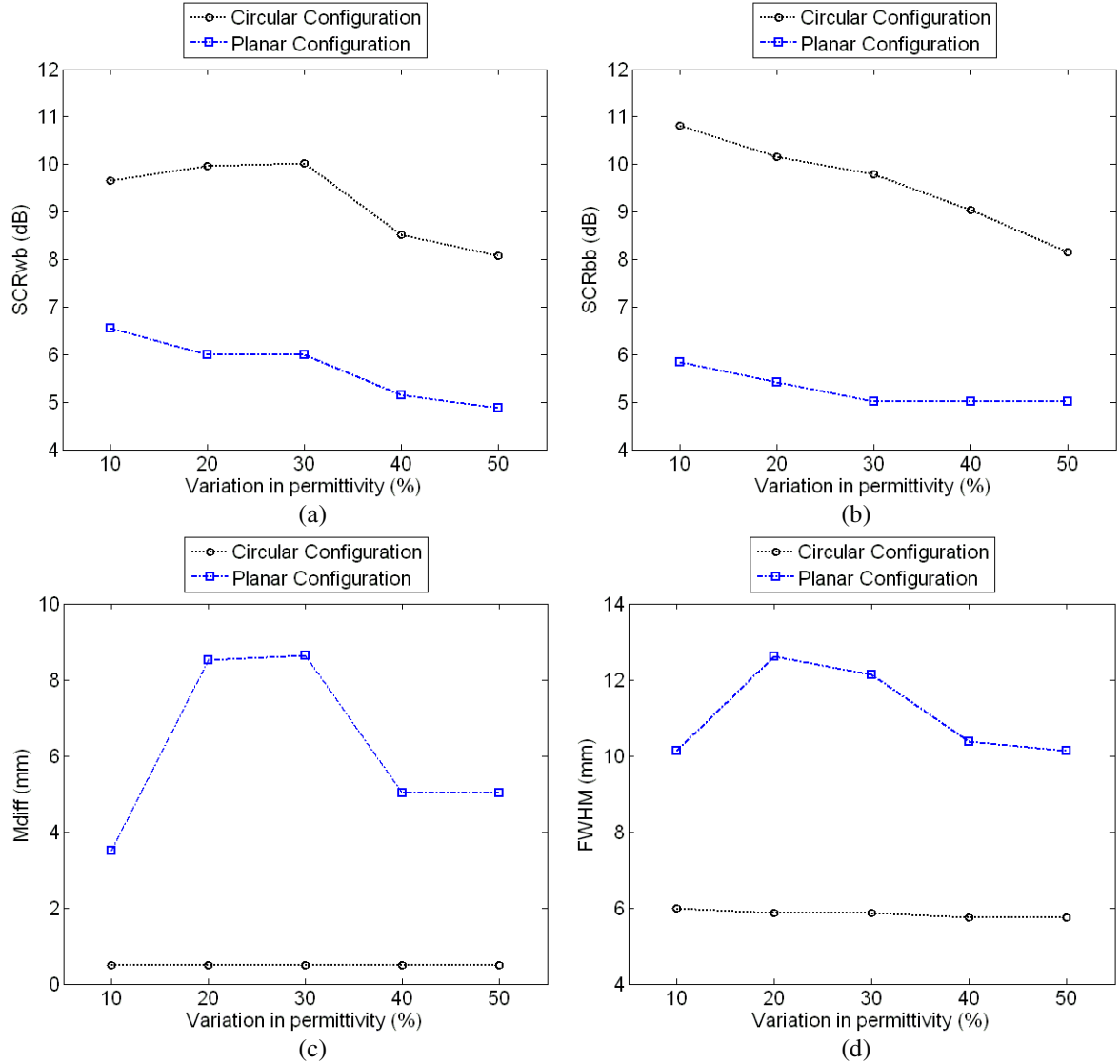


Figure 4.8. Absolute values for four performance metrics varying the relative permittivity from ± 10 to $\pm 50\%$, in steps of 10% , using the circular and planar configurations with 16 monostatic antennas in an FDTD breast model with the presence of a 6 mm tumour.

Furthermore, it is possible to visually and quantitatively detect tumours as small as 2 mm, however, in the interest of brevity, subsequent results and discussion will focus on tumours with diameters of 6 mm. In order to test the robustness of the two antenna configurations with 16 antennas to greater variation in the dielectric properties of normal breast tissue, simulations were run for different percentages of

variation for the relative permittivity values of normal breast tissue, from ± 10 to $\pm 50\%$ in steps of 10%. The results for metrics SCRwb, SCRbb, M_{diff} and FWHM using 16 monostatic antennas and a 6 mm tumour are plotted in Figure 4.8.

Again, the different metrics indicate that the circular configuration performs better than the planar configuration, as observed by higher values for SCRwb and SCRbb, and lower values for M_{diff} and FWHM.

A decrease of SCRwb and SCRbb – Figure 4.8 (a) and (b), respectively – occurs with increased variance in relative permittivity, due to increased clutter in the background of the images. Regarding the M_{diff} and FWHM value metrics – Figure 4.8 (c) and (d), respectively – it is observed that the metrics for the circular configuration remain constant with increasing dielectric variation, whereas the metrics for the planar configuration of antennas display more erratic behaviour.

4.4.2. Optimised Number of Antennas for Varying Tumour Location

In this section, the two antenna configurations are evaluated with a 6 mm tumour placed at different locations in the breast, modelling the different tissues with the same dielectric properties as in Table 4.1. The planar configuration is evaluated with the centre of the tumour located at the following (span,depth) locations: (5,-2.4) cm, (3,-2.4) cm and (1,-2.4) cm. The locations at which the centre of the tumour is placed for the circular configuration are as follows: (3.6,-3.6) cm, (4.1,-3.95) cm and (4.6,-4.3) cm. The performances of the planar and the circular antenna configurations are analysed and compared, using a number of performance metrics, for a varying number of transmitting/receiving antennas (8 to 20) in a total of 390 FDTD simulations with different levels of dielectric heterogeneity, similarly to the first experiment. In this section, the metric SCRbb is replaced by the SMR metric, because of the use of the “ideal” artifact removal. The “ideal” artifact removal is introduced at this stage as it removes the artifact caused by the skin-breast interface more efficiently and the design of an improved artifact removal is out of the scope of this experiment.

For clarity of notation for both antenna configurations, the three locations of the tumour are considered to be at distances 0 , d and $2d$ from the centre of the breast. For the planar and the circular configurations d represents a distance of

approximately 2 cm and 0.61 cm, respectively. The reasoning behind choosing these particular locations for the tumours was to keep an approximately constant distance between the tumour and the skin in each breast model. Consequently, for the planar configuration the tumours could be placed at any place along the span of the breast, and d stands for 2cm. However, for the circular configuration, changing the location of the tumours meant moving them closer to one side of the breast model skin, where d was chosen as 0.61 cm so that the proximity to the skin would have a reduced effect on the results. A diagram showing the locations of the tumours for both configurations is shown in Figure 4.9.

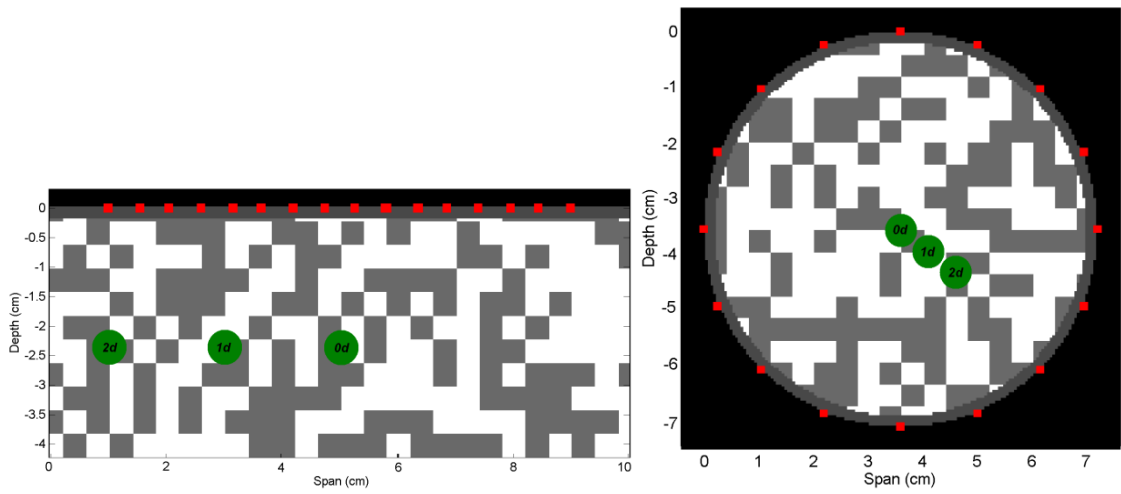


Figure 4.9. The planar and circular antenna configurations are shown on the left and right, respectively. The different locations ($0d$, $1d$ and $2d$) of the 6 mm tumours are shown in green circles. For the planar configuration the tumour is located at: (5,-2.4), (3,-2.4) and (1,-2.4) cm. For the circular configuration the tumour is located at: (3.6,-3.6), (4.1,-3.95) and (4.6,-4.3) cm.

The number of transmit/receive antennas is compared and an optimised number of antennas for improved performance is found through the following performance metrics: SCRwb and SMR. Due to the large amount of data available, the M_{diff} and FWHM metrics are not plotted as they do not vary as much with the varying number of antennas as SCRwb and SMR. The results are shown in Figure 4.10 and Figure 4.11.

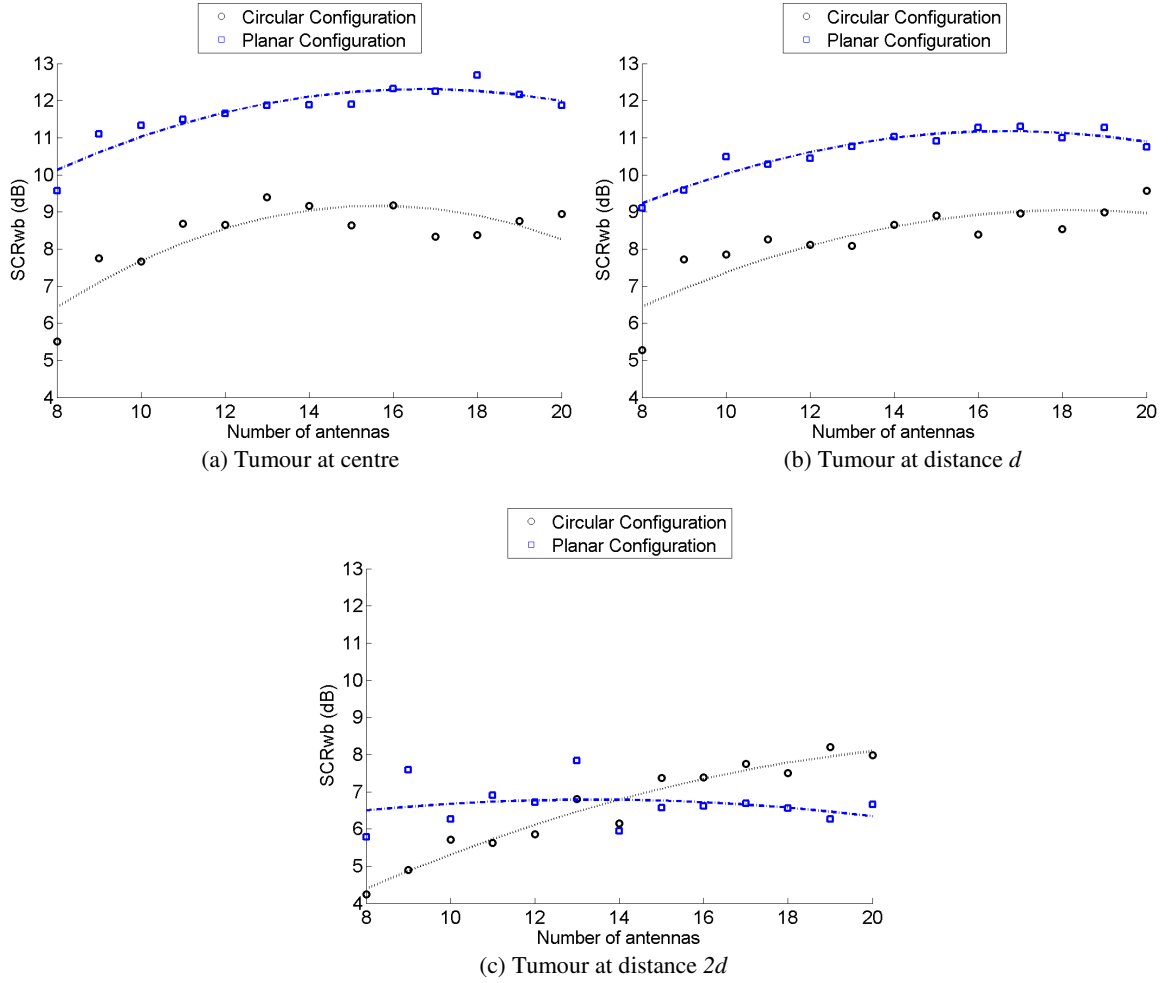


Figure 4.10. SCRwb results versus the number of antennas used for both planar and circular configurations. Quadratic equations have been fitted to the data as follows: (a) $y = -0.005x^2 + 0.211x + 9.355$ (circular) and $y = -0.009x^2 + 0.291x + 11.463$ (planar); (b) $y = -0.009x^2 + 0.356x + 8.216$ (circular) and $y = -0.010x^2 + 0.304x + 11.186$ (planar); (c) $y = -0.018x^2 + 0.666x + 6.483$ (circular) and $y = -0.006x^2 + 0.155x + 10.421$ (planar).

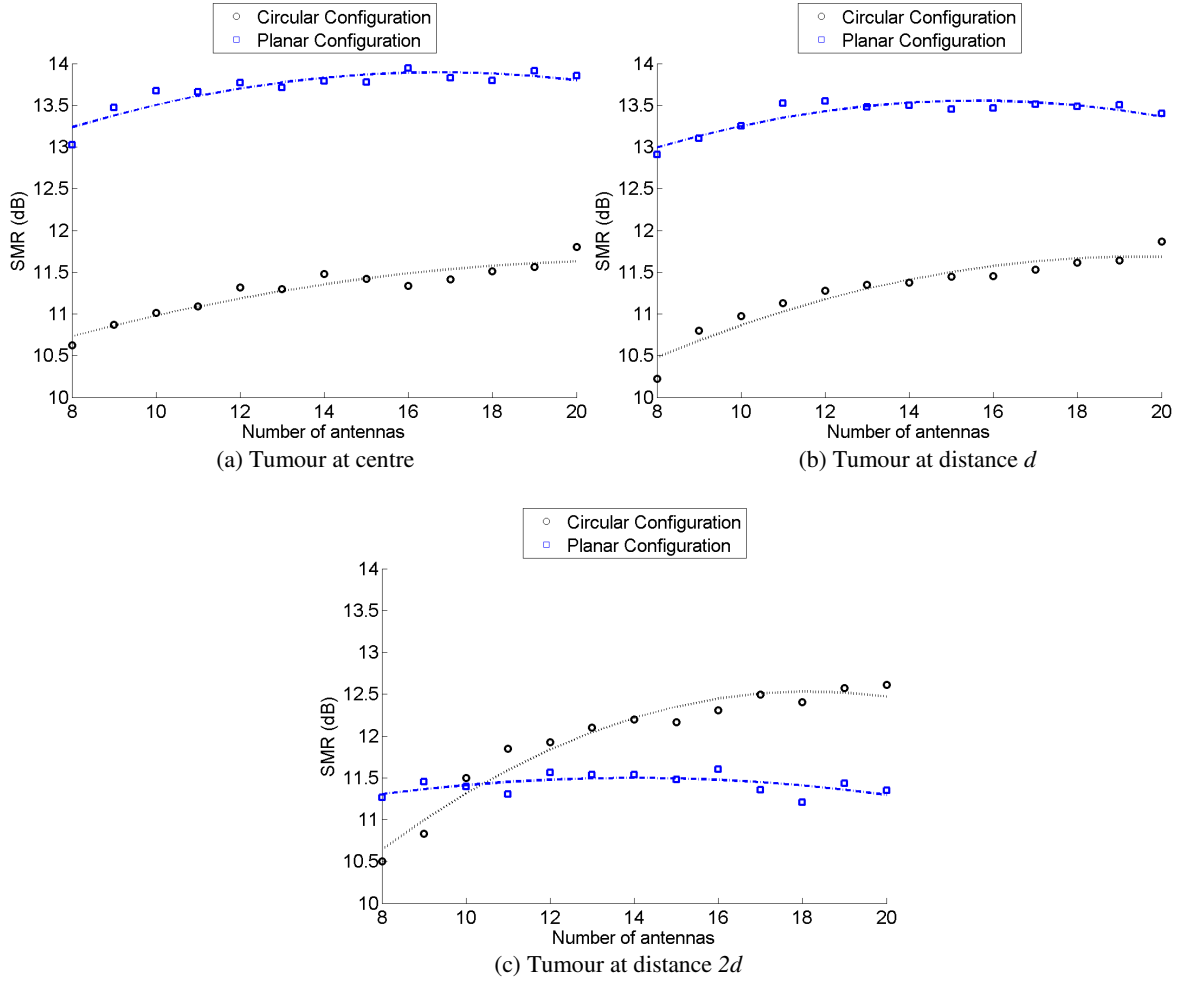


Figure 4.11. SMR results versus the number of antennas used for both planar and circular configurations. Quadratic equations have been fitted to the data as follows: (a) $y = -0.047x^2 + 1.472x - 2.327$ (circular) and $y = -0.029x^2 + 0.968x + 4.248$ (planar); (b) $y = -0.025x^2 + 0.906x + 0.792$ (circular) and $y = -0.026x^2 + 0.869x + 3.948$ (planar); (c) $y = -0.015x^2 + 0.728x - 0.473$ (circular) and $y = -0.010x^2 + 0.269x + 4.990$ (planar).

Both SCRwb and SMR improve with increasing number of antennas for both configurations, with the exception of the planar configuration for the tumour furthest from the centre of the breast. The performance of the circular configuration is virtually independent of the tumour position. Conversely, the performance of the planar configuration tends to worsen when the distance of the tumour from the centre of the breast increases. This behaviour can be explained in terms of the average distances between the known tumour positions and the antennas for each configuration. When the average propagation distance increases, so too do the attenuation and phase effects of the channel, reducing the effectiveness of the beamformer. For the planar configuration, the average propagation distance increases the further the tumour is from the centre of the breast, whereas for the

circular configuration, the average propagation distance does not show the same variation – this is logical since in the circular configuration, if the tumour is further from one antenna, it necessarily means that it moves closer to another one.

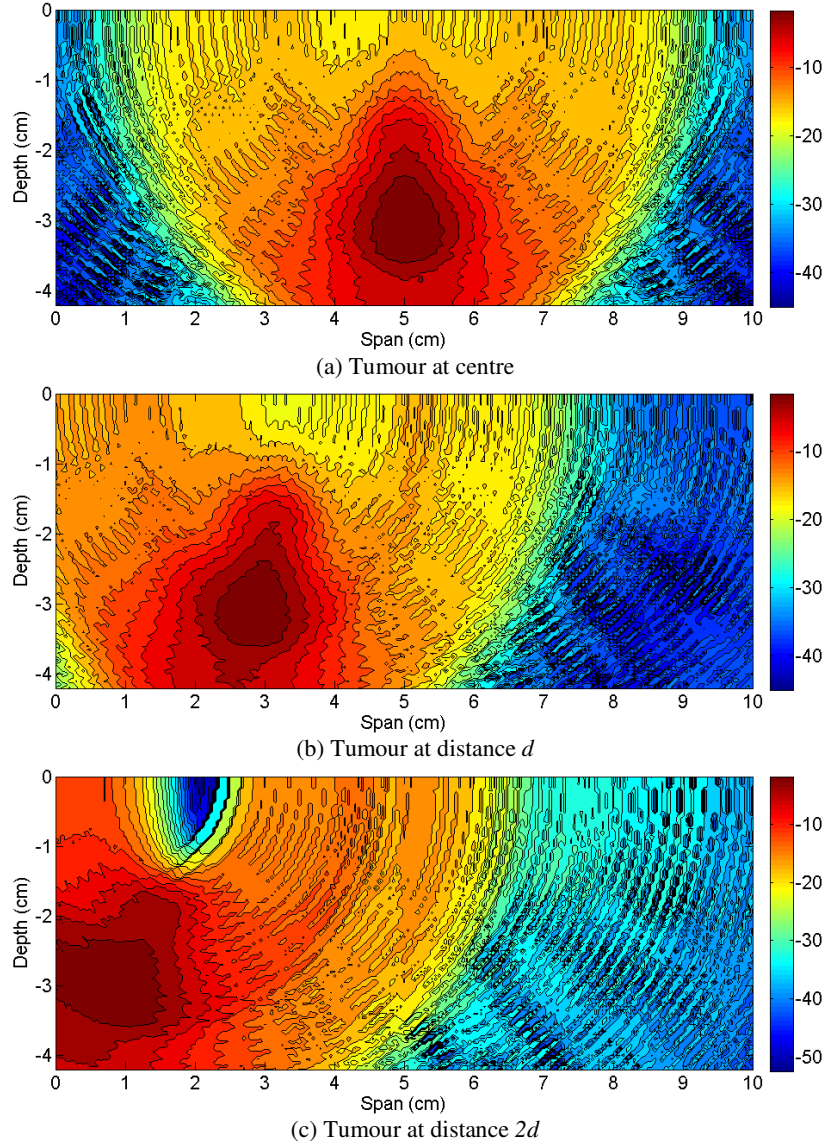


Figure 4.12. Images of backscattered energy (on a dB scale) for the planar antenna configuration using 17 antennas. A 6 mm tumour is centred at: (a) (5,-2.4) cm, (b) (3,-2.4) cm, (c) (1,-2.4) cm. Note different intensity scales for the different plots.

Finally, little improvement is noticed when the number of antennas used exceeds 17, for both antenna configurations, meaning that 17 antennas is found to represent the optimum number of antennas. This finding is largely consistent with the first set of experimental results with tumours of different sizes, where 16 antennas was found to be optimum. The resulting backscattered images for the planar antenna configuration and the circular antenna configuration are shown in

Figure 4.12 and Figure 4.13, respectively. A table comprising the metric results obtained for the optimum number of antennas, 17, is also presented in Table 4.3.

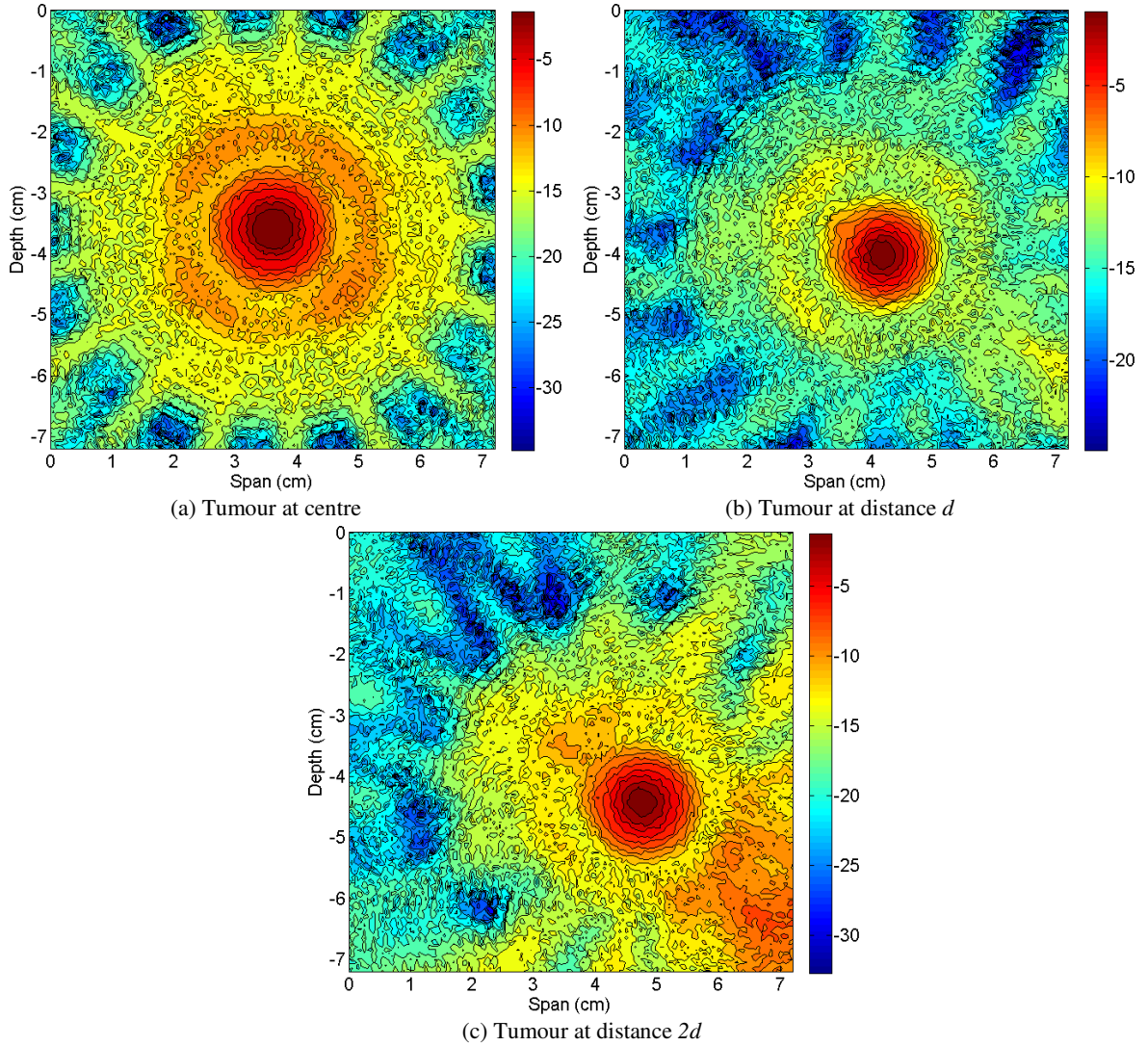


Figure 4.13. Images of backscattered energy (on a dB scale) for the circular antenna configuration using 17 antennas. A 6 mm tumour is centred at: (a) (3.6,-3.6) cm, (b) (4.1,-3.95) cm, (c) (4.6,-4.3) cm. Note different intensity scales for the different plots.

Table 4.3. Results for both antenna configurations with tumours located at three different distances from the centre of the breast (0, d and $2d$) using 17 antennas.

Metric	Antenna Configuration	0	d	$2d$
SCRwb (dB)	Circular	8.34	8.96	7.76
	Planar	12.24	11.31	6.69
SMR (dB)	Circular	11.41	11.53	12.50
	Planar	13.83	13.51	11.36
M_{diff} (mm)	Circular	0.00	1.50	1.80
	Planar	2.69	2.92	4.95
FWHM (mm)	Circular	4.88	4.50	4.25
	Planar	8.00	8.00	9.00

From Figure 4.12 and Figure 4.13, it is also observed that the circular configuration gives a clearer and more isolated position of the tumour when compared to its equivalent results for the planar configuration. This is also shown in the improved FWHM performance offered by the circular configuration, as shown in Table 4.3. Also the centre of the tumour is closer to its actual position for the circular configuration, as reflected in the M_{diff} metric.

As previously observed, using the planar configuration, the tumour appears closer to the skin than it actually is, suggesting that surrounding the breast with the antennas (as in the circular configuration) provides for improved tumour localisation.

Conversely, it is observed that clutter is less significant in the planar images than in the circular images, which is consistent with the results for the SCRwb and the SMR presented in Figure 4.10 and Figure 4.11, respectively. The planar configuration shows better performance in terms of both SCRwb and SMR, except for the $2d$ distance, due to increased average distance between the antennas and the tumour location as previously discussed. However, the circular configuration allows for better localisation of the tumour, indicated by the improved M_{diff} and FWHM results.

Finally, in order to examine the robustness of the two antenna configurations to greater variation in the dielectric properties of normal breast tissue, simulations are performed for different tumour locations, with increasing percentages of variation for dielectric values of normal breast tissue, from $\pm 10\%$ to $\pm 50\%$, in steps of 10% . These simulations all used the optimum number of 17 antennas for both configurations. The results are shown in Figure 4.14.

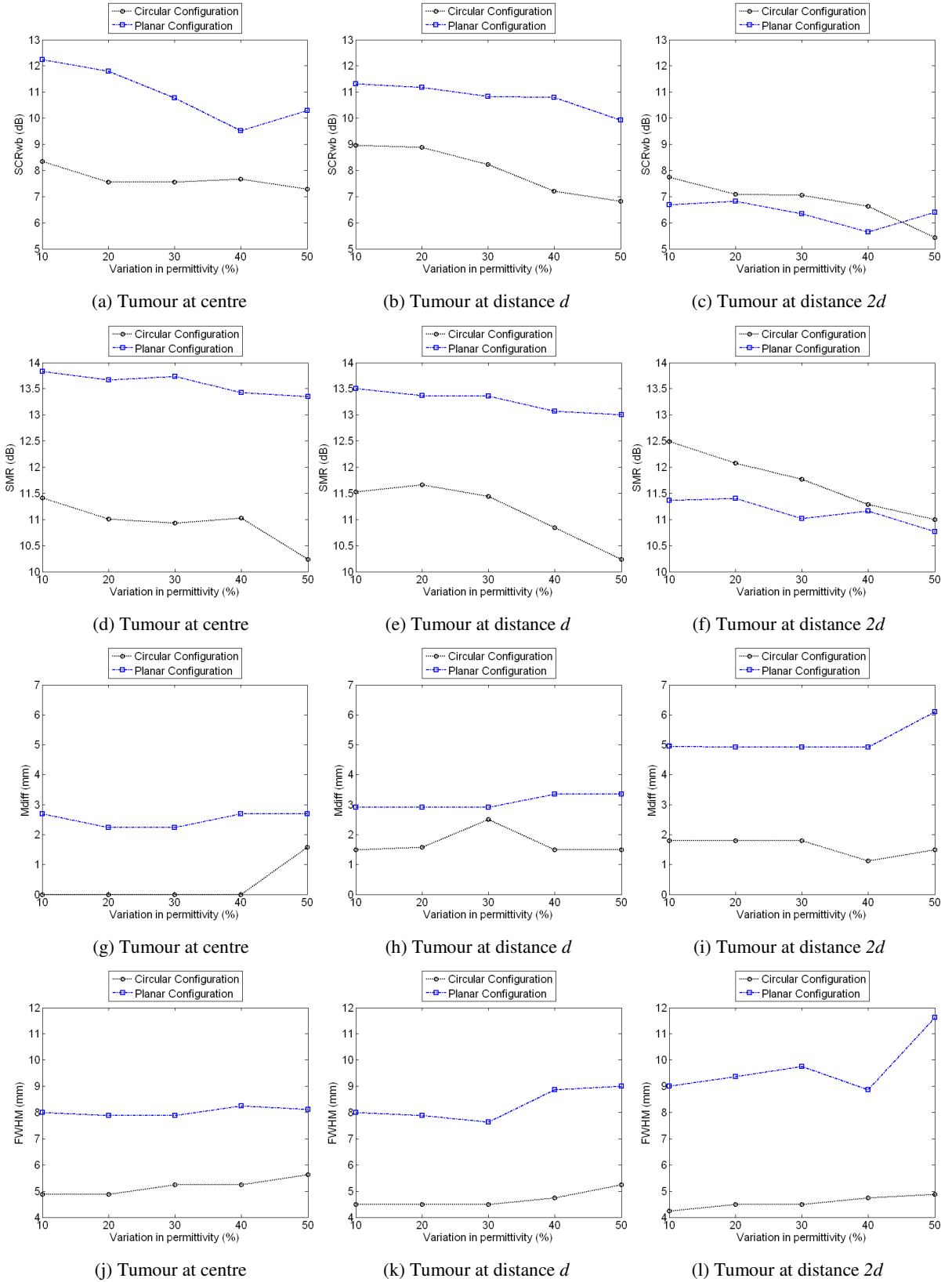


Figure 4.14. Effects of dielectric heterogeneity on both planar and circular antenna configuration. Each row corresponds to the metrics: SCRwb, SMR, M_{diff} and FWHM. Each column corresponds to a tumour located at distance 0, d and $2d$ from the centre of the breast.

Based on Figure 4.14, the SCRwb and the SMR metrics show an overall decrease with the increase of the dielectric variation, for both configurations. The tendency for the planar configuration to outperform the circular configuration when the distance between the tumour and the centre of the breast is 0 or d still holds with variation in dielectric properties. Conversely, the circular configuration outperforms the planar configuration for the simulations in which the tumour is the furthest from the centre of the breast.

Small variations in M_{diff} occur with increasing dielectric variation, suggesting both configurations are relatively robust for this metric. Finally, examining the FWHM, which expresses the physical extent of the tumour response, the circular configuration exhibits very small variation with increasing dielectric variation. This highlights the localisation robustness of the circular system. With the planar configuration, the FWHM of the tumour response remains relatively constant when the tumour is centred in the breast; however, with the increase of the distance between the tumour and the breast centre, the planar system outputs higher values for this metric when the dielectric variation increases, thus reducing the localisation performance of the planar configuration.

4.5. Conclusions

This chapter has examined the effect of antenna configuration on the performance of a UWB system for breast cancer detection. For test purposes, 2D FDTD models of the breast were created, with tumours of different sizes and with tumours at various locations in the breast. Planar and circular antenna configurations were tested through visual inspection and through metrics such as SCRwb, SCRbb, SMR, M_{diff} and FWHM.

Generally it was observed that the circular antenna configuration tended to outperform the planar configuration across the range of metrics, and it exhibits more robust behaviour when increasing dielectric variation is added to the simulations.

For the analysis of tumours of different sizes in the centre of the breast, different numbers of monostatic antennas were tested through visual inspection and four quantitative metrics: SCRwb, SCRbb, M_{diff} and FWHM. The quantitative metrics and visual quality of the images improved when 16 or more antennas were used.

For the analysis of tumours in various locations of the breast, both planar and circular antenna configurations were tested through visual inspection and with four quantitative metrics: SCRwb, SMR, M_{diff} and FWHM. 17 antennas were found to be optimum for both configurations (similar to the optimum number of antennas for experiments with different tumour sizes), as this number offered the best compromise of performance versus efficiency.

While the planar antenna configuration tended to outperform the circular configuration in terms of SCRwb and SMR when the tumour was located close to the centre of the breast, the circular antenna configuration outperformed the planar configuration for tumours further from the centre of the breast.

Furthermore, the circular antenna configuration outperformed the planar configuration across the remaining metrics, suggesting that this configuration is much more effective and robust in terms of tumour localisation, given by M_{diff} and FWHM. The justification for the improved performance of the circular configuration is twofold: Firstly, the greater spatial distribution of the antennas around the entire breast in the circular configuration provides for improved tumour localisation. Secondly, the shorter average propagation distance for signals recorded using the circular antenna configuration results in less attenuated reflections from the tumour.

In conclusion, the circular configuration of antennas shows much more consistent results in terms of tumour identification (SCRwb, SCRbb and SMR), independent of the location of the tumour location within the breast, and generally more precise tumour localisation. It should, however, be mentioned that the planar system allows for the imaging of the underarm, where 50% of the breast cancers occur [28], and therefore the circular-based system should not be used exclusively as it does not cover this region.

The results presented in this chapter compare to those presented in Fear *et al.* [29] in which planar and circular antenna configurations were studied under similar 3D numerical simulations. The results presented in this chapter and in [29] suggest that both planar and circular antenna configurations allow for tumour detection. However in this chapter, detection of tumours with varying sizes and in different locations was examined while that study was not pursued in [29]. Although both studies addressed changing the number of antennas, the results cannot be directly compared since for this chapter the number of antennas was varied within a single 2D array of antennas while in [29] the number of antennas was varied by changing

the number of arrays with antennas, as this was done in a 3D system. Furthermore, it is worth noting that the metrics applied to the antenna configurations in this chapter suggested that the circular configuration outperformed the planar in a number of scenarios.

In the next chapter the classification of tumours will be addressed, using a circular configuration of antennas, since it is important to have robust information on the tumour, independently of its size and location.

Chapter 5

Classification of Breast Tumours in Homogeneous Breast Tissue

5.1.Motivation and Contributions

In addition to detection of tumours based on imaging using UWB radar, there is increasing interest in further characterising tumours, including classification into categories of benign and malignant. However, this task is rendered more difficult by recent findings by Lazebnik *et al.* [26-27] in which it was found that the dielectric contrast between benign and malignant tumour may not be as significant as indicated in previous studies and therefore simple imaging based on backscattered energy may not be effective in this classification task. This problem is addressed in this chapter through investigating the Radar Target Signature (RTS) of dielectric scatterers within the breast, extending work developed by Davis *et al.* [44]. This involves simulating a homogeneous breast with tumours based on GRSs, as described in Chapter 3, which can represent the different stages of growth and development of a tumour through a range of shapes and sizes in a homogeneous breast. A range of features and classifiers are investigated, with a view to developing an effective system for classification of tumours into benign and malignant categories.

The contributions in this chapter are as follows:

- Adding to the three categories of tumour investigated in Davis *et al.* [44], extra granularity is added to the tumour classification process. A fourth tumour model is introduced, the macrolobulated GRS, which represents a benign tumour in a stage of development that may indicate that a tumour is in a pre-malignant stage. The introduction of this tumour model is useful because it is important to classify tumours at the earliest stage of development possible;
- Principal Component Analysis (PCA), Independent Component Analysis

(ICA) and the Discrete Wavelet Transform (DWT) are compared as feature extraction methods for extracting the most significant information from the RTS of the tumours;

- Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Support Vector Machines (SVM) are investigated and compared as methods for classification of the RTS of tumours;
- Eight different multi-stage classification architectures, which are combinations of coarse and/or fine size and shape classifiers, are investigated and evaluated;
- A comprehensive study of classification results are compared between three feature extraction methods, PCA, ICA and DWT, and three classification approaches, LDA, QDA and SVM, using the eight classification architectures.

The results of this work have been published in [50-52, 55, 59], by Conceição *et al.* and McGinley *et al.*

5.2. Tissue and Tumour Simulation

For the experiments reported here, a database of 352 tumours comprising four different sizes and four different shapes was created. The radii of tumours vary between the following discrete values: 2.5, 5, 7.5 and 10 mm. The four types of tumours are spiculated and microlobulated, both malignant types; and macrolobulated and smooth, both benign types.

Both the breast and the tumour tissues are implemented in a 3D FDTD simulation, with a cubic grid resolution of 0.5mm. The FDTD numerical model of the breast is completed with the Debye formulation defined in (2.4), which is reproduced here for completeness:

$$\epsilon_r^*(\omega) = \epsilon_r + \frac{\sigma}{j\omega\epsilon_0} + \frac{\chi_1}{1 + j\omega\tau}$$

The different tissues within the breast are modelled using Debye parameters for malignant tissue and for homogeneous lossy adipose tissue, as established by Lazebnik *et al.* [26-27]. The Debye parameters for malignant tissue are as follows: $\epsilon_\infty = 6.749$, $\Delta\epsilon = 50.09$, $\sigma_s = 0.794 \text{ sm}^{-1}$ and $\tau = 10.50 \text{ ps}$; whereas for

homogeneous lossy adipose tissue they are as follows: $\epsilon_\infty = 3.140$, $\Delta\epsilon = 1.708$, $\sigma_s = 0.036 \text{ sm}^{-1}$ and $\tau = 14.65 \text{ ps}$.

Additionally, the backscattered signals are generated through a Total-Field/Scattered-Field (TF/SF) approach, in which the tumours are completely embedded in the Total-Field (TF) region, similar to [44, 50-51, 59]. For this experiment, the TF/SF region has the following dimensions: the Scattered-Field (SF) is a square geometric prism with square bases measuring 153.5 mm on the side and the height measuring 137.5 mm, while the TF is located at the centre of the SF and is represented by a cube with 40 mm on each side. The origins of the SF and the TF are co-located at (0,0,0) mm [50-51, 59]. The TF/SF region is terminated with a 6 mm Uniaxial Perfectly Matched Layer (UPML), so that the reflections of propagating electric and magnetic fields from the boundaries are absorbed [44, 50-51, 59]. The sampling frequency of the FDTD simulation is 1200 GHz, as before.

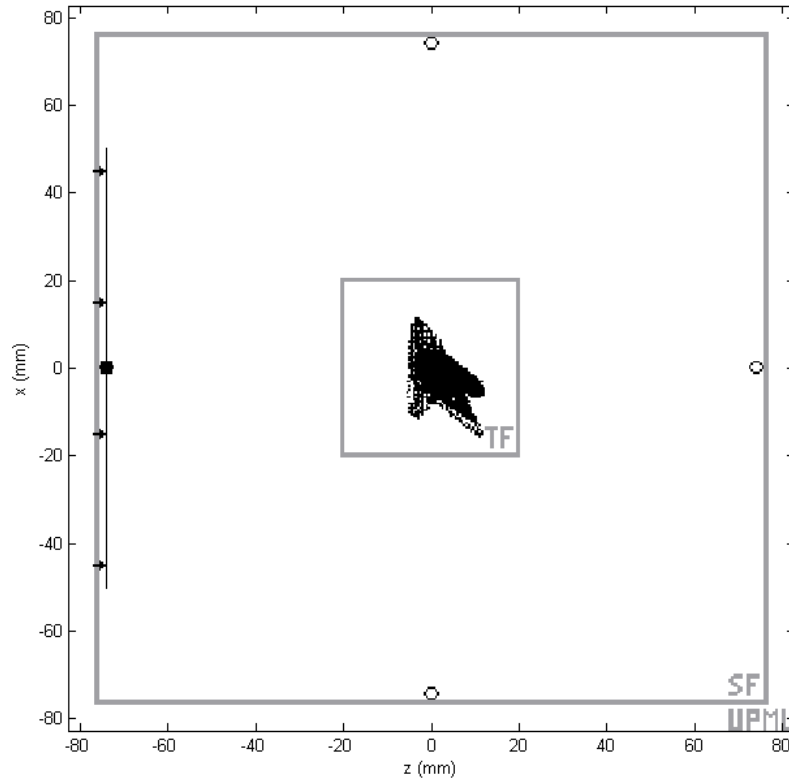


Figure 5.1. Cross-section of the 3D FDTD space lattice partitioned into Total Field (TF), Scattered Field (SF) and UPML regions, for a homogeneous breast model. The target, a spiculated tumour located at the centre of the TF in this example, is illuminated by a pulsed plane wave propagating in the $+z$ direction (represented by a dark line) and backscatter is recorded at the first observer location: (0,0,-74) mm (represented by a filled circle). All four observation points are represented by small circles in the image.

A pulsed plane wave is transmitted towards the target from four equidistant

points, hitting the target at four different angles, equally spaced around the breast circumference. Figure 5.1, adapted from [50-51], shows a representation of the TF/SF grid, with the location of the origin of the first incident plane wave and respective observer point (filled circle) as well as the position of the tumour. All four observation points are represented by small circles.

Each observation point is located in the SF all around the tumour at a constant distance of 74 mm (to match the dimensions of the prototype used in [73, 81, 109]) from the centre of the tumour, which is located at the centre of the TF/SF region. The resulting co-polarised backscatter signals are recorded and analysed from the same four observation points. It must be noted that only co-polarised backscatter is recorded as this was previously found to provide sufficiently high classification performance without the extra computational cost that would involve analysing full polarimetric backscatter [44]. The incident pulse is a modulated Gaussian pulse with centre frequency at 6 GHz and a $1/e$ full temporal width of 160 ps [44]. For the first and third observer locations, (0,0,-74) and (0,0,74), the pulse is linearly polarised in the y and x direction and transmitted in the z direction; while for the second and fourth observer locations, (-74,0,0) and (74,0,0), the pulse is linearly polarised in the y and z direction and transmitted in the x direction. A backscattered signal is recorded without the presence of a tumour and is subtracted from each of the four acquired backscattered signals for each tumour. The resulting signals are finally downsampled from 1200 to 75GHz. Noise was not added to the simulations.

For the purpose of classification, and since eventually the purpose of a classification system is to be coupled with a tumour detection system, such as that described in Chapter 4, the frequency and bandwidth for the input signals used for classification is kept within the same range of the signals used for tumour detection. Although higher frequency pulses would possibly result in higher resolution and may help differentiate between fine differences in the sizes and shapes of tumours, the use of higher frequencies could compromise the detection of tumours which are deeper inside the breast and closer to the chest wall, due to greater attenuation effects.

5.3.Feature Extraction

The backscattered signals related to the four observer locations are firstly

processed through a feature extraction method so that the most relevant features from each signal can be highlighted and forwarded to the classification algorithms. The three feature extraction methods used in this chapter are Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Discrete Wavelet Transform (DWT), which are detailed below.

5.3.1. Principal Component Analysis

Principal Component Analysis (PCA) is a feature extraction method which reduces the dimensionality of multivariate data and reveals simplified structures that are often hidden in the original data set while also disregarding less relevant information such as noise or colinearities in signals [110-111].

This process is accomplished when, by means of a linear algebraic operation, the basis that was used to record the original signals is changed into a new orthonormal basis that allows for a new representation to discriminate the original data set, i.e. when the new data set presents maximal variance. The principal components are ordered by decreasing variance, and furthermore the variance along each principal component provides a measurement of the relative importance of each dimension [111].

For the sake of computational simplicity, the original data should be well represented with the minimum number of principal components, thus creating a problem with the least number of dimensions possible [112]. To obtain the principal components of a matrix X represented by $(m \times n)$, where m is the number of measurements and n is the number of samples, the mean of the sample for each i^{th} measurement is subtracted and finally the basis vectors \mathbf{h}_m , which are the eigenvectors of the covariance matrix $\mathbf{C} = E\{\hat{X}\hat{X}^T\}$, are calculated. The mean-corrected data are represented, for each i^{th} measurement, by its Karhunen-Loève expansion:

$$\mathbf{X} - E\{\mathbf{X}\} = \sum_{m=1}^{N_m} \theta_m \mathbf{h}_m \quad (5.1)$$

in which θ_m represents each basis expansion coefficient and N_m represents the full dimensionality of the problem [44, 111].

5.3.2. Independent Component Analysis

Similarly to PCA, Independent Component Analysis (ICA) is a statistical signal and data analysis method used for multivariate problems in which transformations are applied to the original data so that the new representation of data can capture the essential structure of the data, which can be handled more easily for further analysis. However, for ICA, statistically independent components are extracted through linear transformations from data with a non-Gaussian distribution, whereas for PCA, components are obtained through a linear transformation from data with a Gaussian distribution [113].

To describe ICA, the following notation is defined: the observed data are given by \mathbf{x} which stands for the random vector whose elements are the linear mixtures x_1, \dots, x_n , the independent components are given by \mathbf{s} which is the random vector with elements s_1, \dots, s_n , and n stands for the number of independent components, \mathbf{A} stands for the mixing matrix with the basis vectors a_{ij} , the parameters that relate \mathbf{x} with \mathbf{s} . The ICA model can then be defined as:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (5.2)$$

For the ICA algorithm, only the random vector \mathbf{x} is directly observed, and so both \mathbf{A} and \mathbf{s} have to be estimated. ICA implies that the components s_i are statistically independent and also that they have non-Gaussian distributions. The objective of ICA is to estimate \mathbf{s} and the solution can be represented by:

$$\mathbf{s} = \mathbf{W}\mathbf{x} \quad (5.3)$$

\mathbf{W} is the inverse of \mathbf{A} , which is responsible for the definition of each independent component. Finally, \mathbf{W} can be estimated by iteratively balancing a cost function which either maximises the non-Gaussianity of the calculated independent components or minimises the mutual information, in a process detailed in [113]. There are several algorithms which have been created to efficiently calculate \mathbf{W} ; FastICA is a popular algorithm which is based on a fast fixed-point iteration scheme for maximising non-Gaussianity as a measure of statistical independence for ICA [113-114]. Once \mathbf{W} is estimated, the independent components can be obtained using (5.3).

Comparing ICA to PCA, it is worth noting that ICA has the following limitations:

- The variance of the independent components cannot be determined;

- The order of the independent components cannot be determined [113-114].

5.3.3. Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) can be used as a feature extraction method by producing wavelet coefficients which may be used as discriminant bases for classification methods [115-116].

Wavelets are localised basis functions which are translated and dilated versions of some fixed mother wavelet. The main feature of wavelets is that they are able to highlight localised frequency information about a function or a signal. However, for a successful feature extraction method based on DWT, different wavelet bases have to be evaluated to choose the wavelet which produces optimised classification. For the work presented here, a common library of wavelet filters was examined, which includes well-known wavelets such as: Daubechies, Coiflet, Symlet, Discrete Meyer, Biorthogonal and Reverse Biorthogonal [44, 115-118].

When the DWT is applied to a set of Radar Target Signatures, the wavelet coefficients are obtained by the decomposition low-pass filter and the decomposition high-pass filter. Subsequently, the low-pass band may be split again through a decomposition low-pass filter and a decomposition high-pass filter. It must be noted that for each iteration of the wavelet filters, the number of samples for the next stage is halved through signal decimation. This process continues to a desired number of levels. The final wavelet coefficients, given in a specific frequency subband, are supplied to the subsequent classification methods.

For the results presented in this thesis, the chosen wavelet was *Coiflet 5* [119], as this was the wavelet that allowed for the best performance. *Coiflet 5* was chosen from a comprehensive library of 126 wavelet functions [117, 119] after analysing preliminary classification results with Linear Discriminant Analysis, which is described in detail in the following section. The frequency band that was used corresponds to the wavelet coefficients obtained from the low-pass band obtained after the decomposition of the low-pass filter after two levels of decomposition, as these were found to give the best classification performance compared to other subbands evaluated up to four levels of decompositions. More details are given in Sub-section 5.5.1.

It must be noted that one of Davis *et al.*'s methods for feature extraction, Local

Discriminant Basis (LDB), also seeks to find an orthonormal basis for the data with localised time-frequency characteristics by means of wavelet packets, in specific a third-order *Daubechies* wavelet [44]. The difference between the DWT method used in this thesis and the LDB method is twofold: (i) a different wavelet was used in the thesis, a fifth-order *Coiflet* wavelet, and (ii) also up to four levels of decomposition were obtained to test the full library of wavelet coefficients and a second low-pass decomposition was chosen to allow for the best performance.

5.4. Classification

Three classification methods are used: LDA, QDA and SVM – these are detailed in the following sub-sections. All these classification methods are applied to the processed data using the cross-validation, or leave-one-out, method so that the performance of each classifier is evaluated using a testing set independent from the training set, hence minimising the generalisation error, i.e., the ratio of misclassified samples [120]. The cross-validation method is used as follows. For A -fold cross-validation, the full set of 352 tumours is divided in A subsets, each of which contains one sample of each type and each size of tumours; each subset is then tested against the remaining $(A - 1)$ subsets and, finally, all resulting A sub-classifications are averaged to obtain the overall performance of each classifier.

In addition to these three classifier algorithms, a set of eight multi-stage classification architectures, which categorise the data according to different levels of granularity in size or shape, are also considered in this chapter; these are presented in Sub-section 5.4.4.

5.4.1. Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a technique that allows for discrimination of groups which have multivariate normal Gaussian distributions and have the same covariance matrix. For a discriminant analysis there are dependent variables which represent the classes of the objects and independent variables which are the object features that may describe each class. In case each object is defined by two features, the separator between two groups of objects is a line, otherwise if

objects are defined by three or more features the separator is a hyperplane [121-125].

With the LDA method, a linear combination representing the weighted sum of two or more independent variables defines the discriminant function, which will allow classification [121-125]. The discriminant score for each object, k , in the analysis is a summation of the values obtained by multiplying each independent variable, X_{ik} , by its discriminant weight, W_i . The discriminant Z score for each object is given in the equation below:

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + \dots + W_nX_{nk} \quad (5.4)$$

in which a is the intercept and n is the total number of independent variables. In case there is more than two groups being discriminated at once, a series of classification functions is derived, which depends on the number of each pair of groups [125].

The mean, or centroid, for each group is obtained by averaging the discriminant Z scores for all objects within a group, and it represents the most typical location of an object from a particular group. There are as many centroids as the number of groups being classified at once. The distribution of the discriminant Z scores for each group influence the classification greatly [125].

For LDA, a linear separator is built based on the centroid and the distributions of the discriminant Z scores. Classifiers based on LDA can be applied to the data using the cross-validation method so that the performance of each classifier is evaluated using a testing set, independent from the training set [120].

5.4.2. Quadratic Discriminant Analysis

Quadratic Discriminant Analysis (QDA) is a technique that allows for discrimination of classes which have significantly different class-specific covariance matrices, while the class populations represent multivariate normal Gaussian distributions with the same mean [121-123]. For a discriminant analysis there are dependent variables which represent the classes of the objects and independent variables which are the object features that may describe each class. As a result of having different covariance matrices for each class, the quadratic discriminant function involves both squared and cross-product terms. For a quadratic discriminant function, for a two-class scenario, an individual vector of scores \mathbf{x} is classified as belonging to Group 1 if the following inequality holds:

$$\begin{aligned} & \mu_1'(\Sigma_2^{-1} - \Sigma_1^{-1})x - 2x'(\Sigma_2^{-1}\mu_2 - \Sigma_1^{-1}\mu_1) + (\mu_2'\Sigma_2^{-1}\mu_2 - \mu_1'\Sigma_1^{-1}\mu_1) \\ & \geq \ln \frac{|\Sigma_1|}{|\Sigma_2|} + 2 \ln \frac{|\pi_1|}{|\pi_2|} \end{aligned} \quad (5.5)$$

in which the class means are given by μ_1 and μ_2 , and the covariance matrices are given by Σ_1 and Σ_2 , for Groups 1 and 2, respectively, π_1 and π_2 are the prior probabilities of observing a member of Groups 1 and 2, respectively [124-125].

Generally, QDA offers increased flexibility over LDA at the cost of possibly “overfitting” the training data [122]. Similarly to LDA, classifiers based on QDA can be applied to the data using the cross-validation method so that the performance of each classifier is evaluated using a testing set, independent from the training set [120].

5.4.3. Support Vector Machines

The SVM learning algorithm is typically used as a method to handle nonlinear relations between the input vectors and their corresponding labels by transforming linearly inseparable data to a higher-dimensional space in which they can be more readily separated, usually into two groups [120, 126-130]. For the particular case of the SVM, the input vectors are mapped to a higher-dimension feature space by means of a Kernel (K) [129-130].

The Kernel used for this study is the Radial Basis Function (RBF), which allows for all input vectors to be non-linearly mapped in an infinite-dimension feature space, typically a Hilbert Space. The decision hyperplane can then be obtained in the feature space and is generically given in the following format:

$$w x + b = 0 \quad (5.6)$$

in which w is the normal to the hyperplane, x represents the data and b is the bias.

Knowing that the data can be represented by the inner product $x_i \cdot x_j$ (this is an implication of using an infinite feature space), the equation for the RBF is defined as follows:

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right), \gamma > 0 \quad (5.7)$$

in which γ is the scaling factor of the RBF Kernel [120, 126].

The decision hyperplane is supported by two parallel vectors, one on each side of the hyperplane. Each of these support vectors are at the same distance from the

hyperplane (the “margin”) and each of them delimits either the first or the second labelled class. A classifier will work better when the value for the margin is maximised, so the concept of a soft margin is introduced, as opposed to hard margins, as described in [126, 129-130]. When soft margins are used, it implies that the support vectors are most likely built with supporting samples that represent meaningful samples of the training group, while outlier samples, such as noisy data or unusual data, are ignored for the calculation of the support vectors. If such conditions are met, the learning machine ensures high generalisation [120, 127-128] and therefore will be able to successfully classify an independent testing group. Knowing that the training set is composed of sample-label pairs (x_i, y_i) , in which $i = 1, \dots, l$ represents each sample, x_i represents the input vectors of each sample and y_i represents the respective label, the soft margins can be calculated by following the mathematical optimisation:

$$\min_{w,b,\xi} \left[\frac{1}{2} (\mathbf{w} \cdot \mathbf{w}) + C \sum_{i=1}^l \xi_i \right] \quad (5.8)$$

with the following conditions: $y_i(\mathbf{w} \cdot x_i + b) \geq 1 - \xi_i$, in which the slack variable $\xi_i \geq 0$. For a hard margin the data are scaled so that the margin equals 1, while for a soft margin the margin can be below one as it is given by $1 - \xi_i$. However, this results in the increase of the objective function since the sum of errors, given by $\sum_{i=1}^l \xi_i$, is multiplied by C [120, 130]. The function of C is two-fold: it controls the relative weighting to keep $\mathbf{w} \cdot \mathbf{w}$ small (as the size of the margin is maximised) and it ensures that most samples have a functional margin of at least one [128].

For application of SVM, the data need to be preprocessed so that the SVM classification algorithm can be optimised adequately for the samples, as efficiently as possible. These preprocessing steps are as follows:

- i) Scaling of the training and data set;
- ii) Application of the Kernel function, RBF;
- iii) Application of the Cross-Validation method;
- iv) Optimisation of the RBF parameters.

In the first step, the input vectors (represented by the signals pre-processed with a feature extraction method) for each sample in both training and testing group are scaled to the range $[-1, +1]$. This step is important so that attributes in greater numeric ranges do not dominate those in smaller numeric ranges, and also that the

computational load of the whole algorithm is restricted [120].

For the second step of the SVM algorithm, the RBF is applied.

Another issue that has to be considered is the fact that the combination of (C, γ) for the RBF Kernel has to be tested on the training data through a cross-validation method. This method allows for outlier samples that represent noise or unusual data to be removed, and as a result, some of the outlier supporting samples may be omitted from the final solution.

Finally, the parameters of the chosen Kernel function are adjusted so that the classifier is successful in classifying independent testing groups. For the RBF, the combination of (C, γ) is optimised, C is the penalty parameter of the error term, and as earlier mentioned, γ is the scaling factor of the RBF Kernel. A parameter search, such as the grid-search described in [120], is applied to the data set.

5.4.4. Classification Architectures

Eight different classification architectures are considered, six of which are composed of a number of binary sub-classifiers (and are studied with LDA, QDA and SVM), and the remaining two are composed of direct sub-classifiers that classify four classes at once (and are used with LDA and QDA). SVM is typically used as a binary classifier, and that is the reason why it is not used in the two direct classification architectures in this chapter. The different architectures are defined by the size and shape granularity, i.e. how many categories are classified in each stage (two or four categories), and by the number of stages each size and/or shape classifier is composed of (one or two stages). A coarse shape classifier is used to classify tumours into either malignant or benign tumours, which may, for example, give sufficient information to a doctor in a clinical setting. However, extra granularity in the shape classifier allows further classification of tumours into spiculated, microlobulated (both malignant tumours), and in macrolobulated and smooth (both benign tumours), giving additional clinical information on the development stage of a breast tumour, which may also be useful to a clinician. For instance a macrolobulated shape could potentially be an indicator of pre-malignancy and therefore closer surveillance of the patient may be required.

Each of the classifier architectures will now be described. The first classifier architecture, Coarse-Shape (CS), splits the RTS of the tumours in one stage into two

shape groups: malignant or benign. Similarly, the Fine-Shape (FS) initially classifies the RTS of the tumours into the same shape categories as the CS, but then adds another level of shape granularity by dividing malignant tumours into spiculated and microlobulated tumours and benign tumours into macrolobulated and smooth tumours. The CS and FS architectures can be used with LDA, QDA and SVM classifiers and are both shown in Figure 5.2.

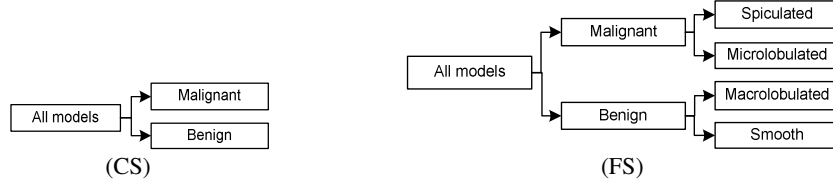


Figure 5.2. Classification architecture in which only shape classifications are applied: CS and FS.

The Coarse-Size-Coarse-Shape (CSCS) splits the RTS of the tumours in one stage into two size groups (the first group contains 2.5 and 5 mm tumours and the second group contains 7.5 and 10 mm tumours), before further classifying the tumours into either benign or malignant. Similarly, the Coarse-Size-Fine-Shape (CSFS) initially classifies the RTS of the tumours into the same size and shape categories as the CSCS, but then adds another level of shape granularity by dividing malignant tumours into spiculated and microlobulated tumours, and benign tumours into macrolobulated and smooth tumours in a second stage of shape classification. The CSCS and CSFS architectures can be used with LDA, QDA and SVM classifiers and are both shown in Figure 5.3.

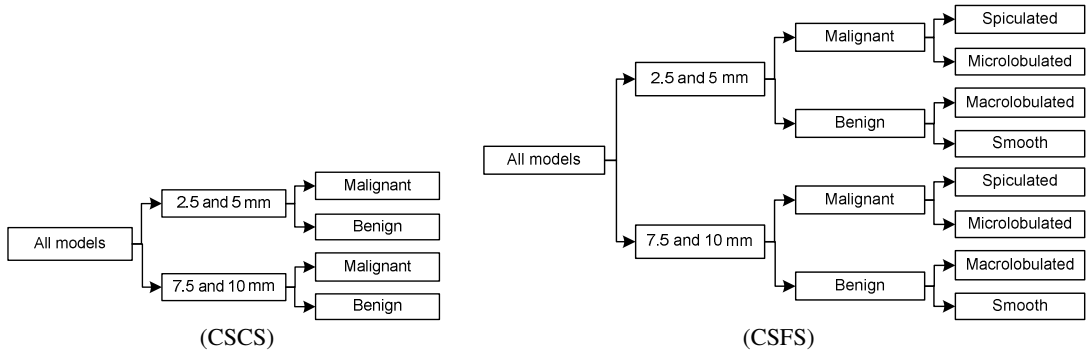


Figure 5.3. Classification architectures in which a 1-stage coarse size classification is applied before shape classification: CSCS and CSFS.

The Fine-Size-Coarse-Shape (FSCS) and Fine-Size-Fine-Shape (FSFS) classification architectures further classify the RTS of the tumours into four

subcategories of size (2.5mm, 5mm, 7.5mm and 10mm) in two stages. The FSCS then divides them into two categories of shape, benign and malignant, while the FSFS classifies them into four shape categories: spiculated, microlobulated, macrolobulated and smooth in two stages. The FSCS and FSFS classifiers can be used with LDA, QDA and SVM classifiers and are shown in Figure 5.4.

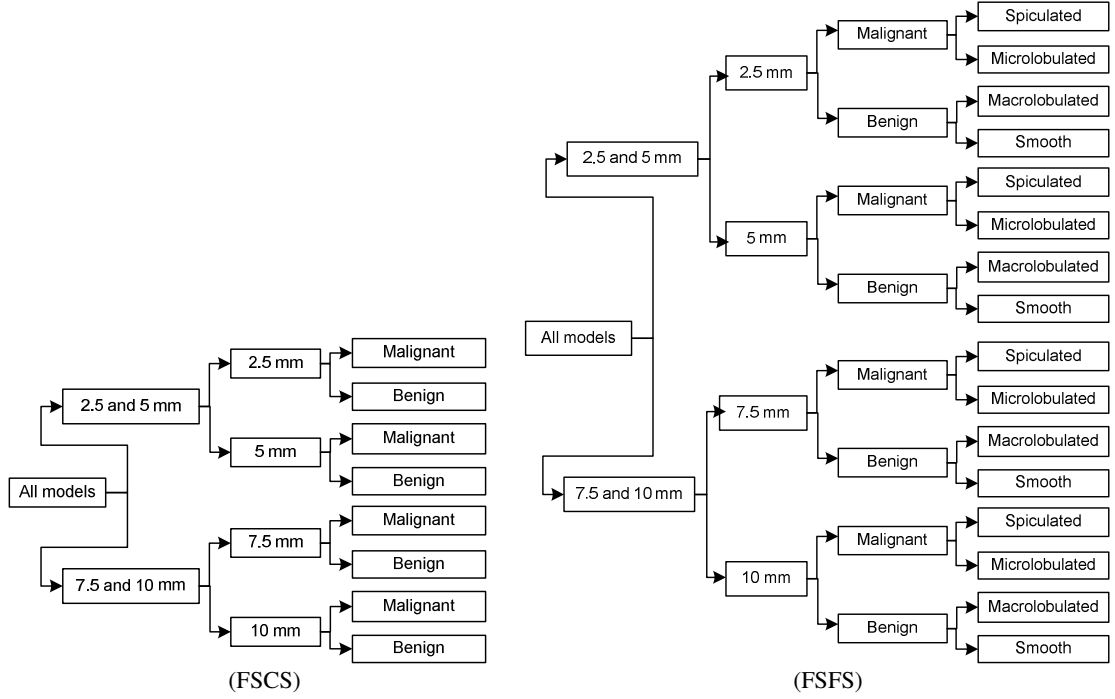


Figure 5.4. Classification architectures in which a 2-stage fine size classification is applied before shape classification: FSCS and FSFS.

The Direct-Fine-Shape (DFS) classification architecture performs the same function as the FS classifier, but divides the RTS of the tumours into four shape categories in one stage. The Direct-Fine-Size-Fine-Shape (DFSFS) performs the same function as the FSFS classifier, but firstly divides the RTS into four size categories, and then classifies the RTS of the tumours into the four shape categories in a second stage. The architectures of the DFS and the DFSFS classifiers can be used with LDA and QDA classifiers and are shown in Figure 5.5.

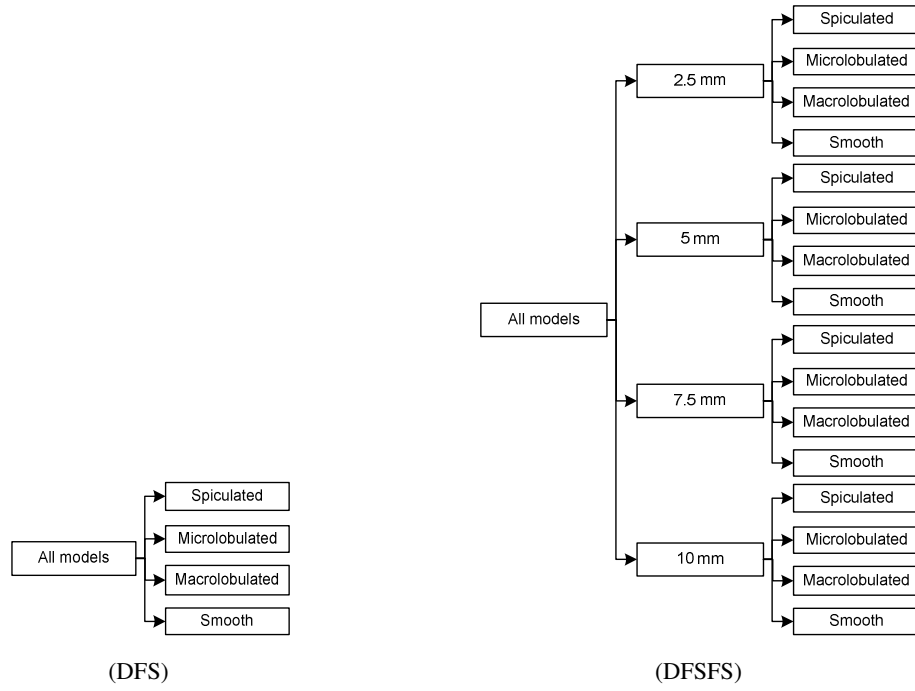


Figure 5.5. The classification architectures in which a direct fine size and/or shape classifications are applied: DFS and DFSFS.

The experiments are divided into four parts, which are described in more detail in Section 5.5:

- Firstly, different feature extraction methods are used to extract the most significant bases from the RTS of the tumours: PCA, ICA and DWT (with *Coiflet 5* wavelet). In particular, a dimensionality reduction of PCA is accomplished as a pre-processing step so that the more representative bases of the data are selected. The method to optimise feature extraction by means of DWT is briefly detailed.
- Secondly, the backscattered signals obtained in four observation points are compared in terms of LDA and QDA when using the three feature extraction methods.
- Thirdly, the SVM classifier is tested on the data obtained after the application of the three feature extraction methods.
- Finally, the three feature extraction methods are evaluated and their effects on three classifiers are compared in a single comprehensive study.

5.5. Results and Discussion

5.5.1. Comparison of PCA, ICA and DWT

PCA, ICA and DWT are used as feature extraction methods to extract the most important features that will allow the determination of the RTS of the different tumours, which will then be classified. In this sub-section the issue of dimensionality and how the feature extraction methods are applied are addressed.

Firstly, the dimensionality reduction of PCA is investigated. For that purpose, performance is obtained for both a coarse size (partial size classifier in CSCS and CSFS) and a coarse shape (CS) classifiers for varying numbers of PCA components, from 2 to 100, using both LDA and QDA approaches. For this part of the experiment the backscattered results from the four observation points are used. These graphs are shown in Figure 5.6, in which the accuracy of LDA (black line) and QDA (blue line) classifiers versus the number of principal components is plotted. The performance results for the coarse size classifier are shown on the left and for the coarse shape classifier on the right.

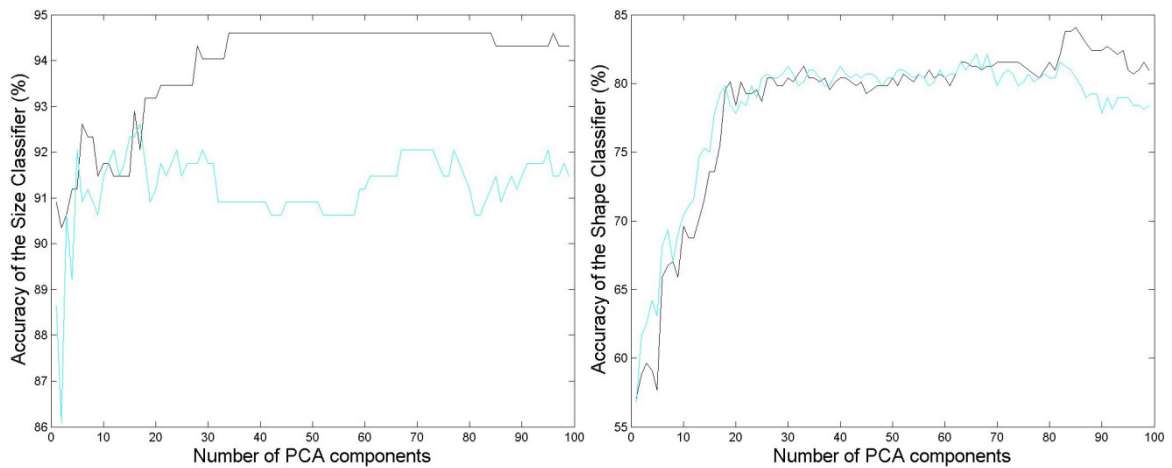


Figure 5.6. Accuracy of LDA (black line) and QDA (blue line) classifiers versus the number of principal components used. The accuracy results for the coarse size classifier are shown on the left and the accuracy results for the coarse shape classifier are shown on the right. Note that the axes are different for the two graphs.

From Figure 5.6, it can be seen that a dimensionality in the range of 25-30 components appears to be a reasonable compromise between computational complexity and performance, as the classification performances start to saturate at

this level. Therefore, the use of 30 principal components offers a good compromise between classification accuracy and computational time, which had also been found in [44].

Regarding ICA, the FastICA algorithm [114] outputs 16 independent components to represent each signal.

For the DWT method, the wavelet function in use is *Coiflet 5*, and the wavelet coefficients from the low-pass subband after the decomposition of the low-pass filter with two iterations represent the best features for classification. As the DWT algorithm is applied twice, the original length of the signal is downsampled to half in two instances. Given that the original RTS signal used for analysis is 100 samples in length, the final number of samples used for classification is 25. A simple block diagram of the DWT decomposition “tree” with the relevant sub-band highlighted is presented in Figure 5.7.

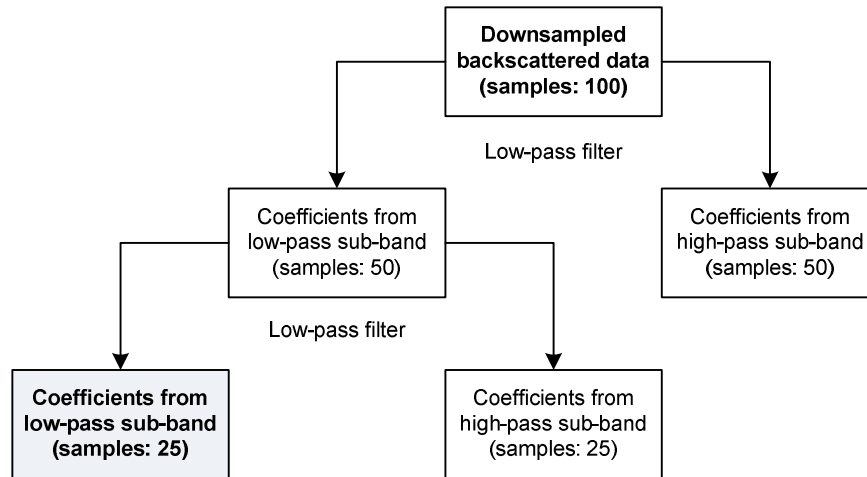


Figure 5.7. Simple block diagram of the DWT decomposition “tree” with the relevant sub-band highlighted.

5.5.2. Comparison between LDA and QDA Classifiers

In this sub-section the eight binary and direct classification architectures, described in Section 5.4.4, are compared using two specific classifiers, LDA and QDA, to assess the shape of tumour models. The performance for all classification architectures are shown for all combinations, between the LDA and QDA classifiers and the three feature extraction methods, in Table 5.1. The results are shown in three columns which represent:

- The partial accuracy for the size classification;
- The partial accuracy for the shape classification;
- The overall accuracy for the size-then-shape classification.

The accuracy of the partial size and the partial shape classifiers is expressed in terms of the proportion of tumours correctly identified in terms of size or shape, respectively, while considering those attributes in isolation. The overall accuracy for the size-then-shape classifier is calculated by multiplying the partial accuracies for the size and shape classifiers and represents the percentage of tumours correctly classified in terms of both size and shape. Finally, the best performance for each stage of the classification architectures is highlighted in blue. Note that several of the architectures are single-stage, and therefore “partial classification” into particular categories is not applicable in these cases (e.g. partial size classification is not applicable for the CS classifier).

A first observation from Table 5.1 is that applying a size classifier to the whole dataset of tumours offers higher accuracy than the shape classifier; this is consistent with the results in Figure 5.6. In absolute terms, the best performance of the coarse size classifier is 13.64% higher than the best performance of the coarse shape classifier (94.32% versus 80.68%). As for the best fine size classifier and fine shape classifier, there is an improvement of 28.69% (85.51% versus 56.82%). This provides a measure of justification for the use a cascade of a size classifier followed by a shape classifier to provide a more accurate shape classification, as reflected in the CSCS, CSFS, FSCS, FSFS and DFSFS architectures.

It can also be observed from Table 5.1 that the PCA and the DWT feature extraction methods allow for more accurate classifications with both LDA and QDA than those obtained through the components selected with ICA. By averaging all classification accuracies obtained for both partial size and shape classifiers, individually for each feature extraction method, it was found that while PCA and DWT-based classifications perform similarly with an average between 74% and 75%, the classifications based on ICA perform with an average of approximately 68%. Between PCA and DWT-based features, they both result in similar LDA and QDA performance, with DWT-based classification generally outperforming PCA-based classifications. Examining Table 5.1, it must be noted that there is a form of “error propagation” when a partial coarse size (or shape) classifier is extended to a

fine size (or shape) classifier, as any misclassified tumours in the first stage of the size (or shape) classification are necessarily misclassified in the second stage of size (or shape) classification. This error propagation results in an absolute performance decrease of 8.81% for the best partial size classification and between 17.90% and 23.86% for the best partial shape classification.

Table 5.1. Accuracy for size classifier, subsequent shape classifiers and overall size-then-shape classifier combining feature extractions (PCA, ICA and DWT) with LDA and QDA classifiers for eight different classification architectures. The best accuracy for each stage of the classification architecture is highlighted.

Architectures of classifiers	Feature extraction methods	Partial size classifier (%)		Partial shape classifier (%)		Size-then-shape classifier (%)	
		LDA	QDA	LDA	QDA	LDA	QDA
Coarse-Shape	PCA	N/A		79.83	80.68	N/A	
	ICA			73.58	75.00		
	DWT			79.54	80.11		
Fine-Shape	PCA	N/A		54.83	56.82	N/A	
	ICA			41.19	45.17		
	DWT			54.26	53.98		
Coarse-Size-Coarse-Shape	PCA	94.03	92.04	85.23	79.26	80.14	72.96
	ICA	91.48	92.33	76.14	77.27	69.65	71.34
	DWT	94.32	92.04	85.23	79.54	80.38	73.22
Coarse-Size-Fine-Shape	PCA	94.03	92.04	65.91	56.82	61.98	52.30
	ICA	91.48	92.33	53.12	53.98	48.60	49.84
	DWT	94.32	92.04	65.34	59.09	61.63	54.39
Fine-Size-Coarse-Shape	PCA	84.37	74.15	85.51	83.81	72.15	62.14
	ICA	79.54	72.16	79.83	73.86	63.50	53.30
	DWT	85.51	73.86	86.08	82.67	73.61	61.06
Fine-Size-Fine-Shape	PCA	84.37	74.15	68.18	63.64	57.53	47.18
	ICA	79.54	72.16	57.95	52.84	46.10	38.13
	DWT	85.51	73.86	68.18	58.81	58.30	43.44
Direct-Fine-Shape	PCA	N/A		54.26	59.37	N/A	
	ICA			42.90	38.07		
	DWT			55.40	53.12		
Direct-Fine-Size-Fine-Shape	PCA	81.53	74.71	59.94	61.65	48.87	46.06
	ICA	72.44	71.59	51.42	58.52	37.25	41.90
	DWT	82.39	75.57	63.64	56.53	52.43	42.72

Results obtained with PCA and DWT are now discussed. Examining the CS/FS, CSCS/CSFS and FSCS/FSFS architectures, the results in Table 5.1 show that the accuracy for the partial shape classification is considerably higher when a previous size classifier is used for both LDA and QDA, with an exception for the shape classifier in CSCS when using QDA as the corresponding performance in CS is higher. Even though the accuracy of the fine size classifier (in FSCS/FSFS) is lower than the coarse classifier (in CSCS/CSFS), the higher granularity of the fine size

classifier allows for better performance in the subsequent shape classifiers.

In summary, examining the overall performance of the classification architectures, there are several findings:

- When considering the overall accuracy of a size-then-shape cascade classifier, the accuracy of the cascade classifiers worsens with increasing number of stages in the partial size and/or shape classifications, i.e. when fine-size and/or fine-shape classifiers are used. This may not be unexpected, since fine classification is more challenging than coarse classification.
- FS and DFS are the least accurate shape classifiers, leading to the conclusion that a shape classifier is more accurate when preceded by a size classifier.
- There is little performance difference between the FSFS and DFSFS architectures (the architectures with highest granularity in both size and shape), for both PCA and DWT feature extraction methods. FSFS and DFSFS performed better with LDA instead of QDA. However it must also be noted that DFSFS results do not deteriorate to the same extent as FSFS when using the alternative discriminant analysis method, QDA, suggesting that FSFS is more robust.
- The best classifier to simply detect whether a tumour is malignant or benign is FSCS for both QDA and LDA, in which a fine-size classifier is used.
- To obtain a clearer discrimination of the shape of tumours beyond simply malignant and benign (spiculated, microlobulated, macrolobulated or smooth), the FSFS performs best.

Also, for the results based on the PCA and DWT feature extraction methods, there is no single “better” classification method. However it is observed that generally LDA outperforms QDA in the used architectures.

5.5.3. Classification Using SVM

In this chapter a novel classification scheme based on SVM is applied to the six binary-based classification architectures described in Section 5.4.4 (CS, FS, CSCS, CSFS, FSCS and FSFS). The performance for the SVM-based classification architectures depends on optimal values for parameters γ and C , as described in 5.4.3. The downsampled UWB backscattered signals are processed through the same three feature extraction methods: PCA, ICA and DWT, and the SVM parameters are

optimised for best classification performance with each feature, through a grid-search procedure.

For the results presented here, the optimal SVM parameters are as follows: for PCA $\gamma=2^3$ and $C=2^8$, for ICA $\gamma=2^5$ and $C=2^{25}$ and for DWT $\gamma=2^7$ and $C=2^{27}$. Results for the different binary-based classification architectures are shown in Table 5.2, in which the best accuracy for each stage of the classification architectures is highlighted. The results are shown in three columns in the same fashion as for Table 5.1.

Table 5.2. Accuracy for size and subsequent shape classifiers and overall size-then-shape classifier using SVM binary classifier for six different classification architectures. The best accuracy for each stage of the classification architecture is highlighted.

Architectures of classifiers	Feature extraction methods	Partial size classifier (%)	Partial shape classifier (%)	Size-then-shape classifier (%)
Coarse-Shape	PCA	N/A	89.20	N/A
	ICA		84.66	
	DWT		86.08	
Fine-Shape	PCA	N/A	72.73	N/A
	ICA		66.19	
	DWT		67.33	
Coarse-Size-Coarse-Shape	PCA	94.89	90.62	85.99
	ICA	94.32	87.21	82.26
	DWT	94.60	91.19	86.27
Coarse-Size-Fine-Shape	PCA	94.89	75.00	71.16
	ICA	94.32	70.17	66.18
	DWT	94.60	74.43	70.41
Fine-Size-Coarse-Shape	PCA	86.93	90.34	78.53
	ICA	79.83	86.93	69.40
	DWT	84.38	90.91	76.70
Fine-Size-Fine-Shape	PCA	86.93	75.28	65.44
	ICA	79.83	69.60	55.56
	DWT	84.38	75.00	63.28

After a first analysis of Table 5.2 it is reasonable to assume that PCA as a feature extraction method allows for better classification using SVM in most cases when compared to the alternative features. By averaging all classification accuracy percentages obtained for both partial size and shape classifiers, individually for each feature extraction method, it is confirmed that while PCA and DWT-based classifications perform similarly with an average of approximately 86% and 84%, respectively, the classifications based on ICA perform with an average of only 81%, and so the ICA feature extraction method is the one that gives poorer discriminative

information for the SVM classifier, similar to the results for LDA and QDA reported in Table 5.1.

Similarly to Table 5.1, there is an “error propagation” in Table 5.2 when a partial coarse size (or shape) classifier is extended to a fine classifier. This error propagation results in an absolute performance decrease of 7.96% for the best partial size classification and between 15.63% and 16.47% for the best partial shape classification.

Examining the six architectures in Table 5.2, it can be observed that applying a size classifier, both coarse and fine, to the whole group of 352 tumours, results in better performance than applying a shape classifier in isolation, also for both coarse and fine classifiers. In absolute terms, the performance of the best coarse size classifier is 5.69% higher than the best coarse shape classifier, and the best fine size classifier is 14.20% higher than the best fine shape classifier when applied directly to the whole data set. Given these results it is sensible to investigate the performance of a shape classifier when a size classifier is previously applied, like in architectures CSCS, CSFS (Figure 5.3), FSCS and FSFS (Figure 5.4).

It is observed that the partial shape classifier has slightly higher performance when there is a previous size classifier compared to when there is no previous size classification: the partial coarse shape classifier is highest in CSCS, 91.19% and the partial fine shape classifier is highest in FSFS, 75.28%. However, it must be noted that the performance of the partial shape classifier does not vary significantly whether there is a previous size classifier (and therefore a cascade classifier) or not; in absolute terms, the performance of the coarse shape classifier in CSCS is 1.42% higher than in CS (89.20% versus 90.62%) and the performance of the fine shape classifier in FSFS is 2.55% higher than in FS (72.73% versus 75.28%).

In terms of the overall performance of the classifiers it is observed that the more stages of partial size and/or shape classifiers the lower the overall performance. This might be expected, as the greater the level of granularity in the classification, the more challenging the problem. Conversely, the fewer stages of partial size and/or shape classifiers the higher the overall performance. The overall performance drops by 23.76%, in absolute terms, from the architectures CS to FSFS. The performance of the classifiers decreases with the increase of granularity for two specific reasons: Firstly, for fine classification (e.g., differentiating between smooth and

macrolobulated), the RTS of the tumours are quite similar, so classification is much more difficult and misclassifications are much more likely to occur. Secondly, when classifiers are grouped in architectures such as the ones used in this study, errors can propagate through the multi-stage classifier. For instance, a microlobulated tumour which is first classified as benign will never be classified correctly in a fine shape classifier (as it will automatically be misclassified as a smooth or a macrolobulated tumour).

For analysis purposes, errors in classification in terms of both size and shape were recorded at each stage of the classification architectures. In general it was observed that the number of tumours misclassified for one class was very similar to the number of tumours misclassified for the second class (for binary classification as used here for each individual stage). However, there were two exceptions:

- The fine shape classifier misclassified several smooth tumours as macrolobulated and vice-versa despite the type of size pre-classification. This is due to the similarity between these two types of benign tumours.
- There was a significant number of spiculated tumours misclassified as microlobulated and vice-versa for larger tumours (with 7.5 and 10 mm radius). This can be explained by the fact that the spicules of the spiculated tumours are always the same length (independent of the tumour radius). Therefore, in smaller tumours models, the spicules extend further beyond the surface of the tumour compared to larger tumours. In larger malignant tumours, the spicules may not influence the RTS of the tumours as much as for smaller spiculated tumours, and therefore misclassifications between larger spiculated and microlobulated tumours are more likely to occur.

5.5.4. Overall Evaluation of Feature Extraction and Classification

The results discussed in Sections 5.5.2 and 5.5.3 are combined in Table 5.3 for direct comparison on the database of 352 tumour models.

Table 5.3. Summary table for the classification performance with PCA, ICA and DWT as feature extraction methods, and LDA, QDA and SVM classifiers for a range of different classification architectures. The best accuracy for each stage of the classification architecture is highlighted.

Architectures of classifiers	Feature extraction methods	Partial size classifier (%)			Partial shape classifier (%)			Size-then-shape classifier (%)		
		LDA	QDA	SVM	LDA	QDA	SVM	LDA	QDA	SVM
Coarse-Size-Coarse-Shape	PCA	94.03	92.04	94.89	85.23	79.26	90.62	80.14	72.96	85.99
	ICA	91.48	92.33	94.32	76.14	77.27	87.21	69.65	71.34	82.26
	DWT	94.32	92.04	94.60	85.23	79.54	91.19	80.38	73.22	86.27
Coarse-Size-Fine-Shape	PCA	94.03	92.04	94.89	65.91	56.82	75.00	61.98	52.30	71.16
	ICA	91.48	92.33	94.32	53.12	53.98	70.17	48.60	49.84	66.18
	DWT	94.32	92.04	94.60	65.34	59.09	74.43	61.63	54.39	70.41
Fine-Size-Coarse-Shape	PCA	84.37	74.15	86.93	85.51	83.81	90.34	72.15	62.14	78.53
	ICA	79.54	72.16	79.83	79.83	73.86	86.93	63.50	53.30	69.40
	DWT	85.51	73.86	84.38	86.08	82.67	90.91	73.61	61.06	76.70
Fine-Size-Fine-Shape	PCA	84.37	74.15	86.93	68.18	63.64	75.28	57.53	47.18	65.44
	ICA	79.54	72.16	79.83	57.95	52.84	69.60	46.10	38.13	55.56
	DWT	85.51	73.86	84.38	68.18	58.81	75.00	58.30	43.44	63.28

Based on the data presented in Table 5.3, the overall average results between each combination of feature extraction method (PCA, ICA and DWT) and classification methods (LDA, QDA and SVM), considering both partial size and shape sub-classifiers, are calculated and presented in Table 5.4. The accuracy for the size-then-shape classifier is not considered in this calculation since it is a simple multiplication of the partial size and shape sub-classifiers.

Table 5.4. Averaged results for each combination of feature extraction and classification methods, for partial size and shape sub-classifiers. For each classification method, the feature extraction method showing the best average accuracy is highlighted.

Feature extraction methods	Classification methods		
	LDA	QDA	SVM
PCA	82.7	77.0	86.9
ICA	76.1	73.4	82.8
DWT	83.1	76.5	86.2

In general, the best feature extraction method for use with the LDA classifier (across the range of classifier architectures) was found to be DWT, with a performance average of approximately 83.1%. The DWT outperformed both PCA and ICA by approximately 0.4% and 7% on average, respectively. Similarly, in the case of QDA, the optimum feature extraction method was PCA, with a performance average of 77.0%, which was shown to provide an approximate average

improvement of 3.6% and 0.5% over ICA and DWT, respectively. Finally, when the SVM classifier is considered, the best feature extraction method was found to be PCA, with a performance average of 86.9%, which outperformed both ICA and DWT by approximately 4.1% and 0.7%, respectively.

In terms of the best combination of feature extraction method and classifier for size classification, both coarse and fine classification performed best with PCA combined with SVM, achieving a performance of 94.89% and 86.93%, respectively (Table 5.3). Regarding shape classification, the best combination of feature extraction method and classifier for simply determining whether a tumour was benign or malignant was SVM with the DWT, which provided the best results at 91.19% (Table 5.3). Finally, if extra shape granularity is required, the optimum configuration is SVM used with the PCA feature extraction method, yielding an overall performance of 75.28% (Table 5.3).

5.6. Conclusions

This chapter of the thesis has described experiments conducted on the analysis of UWB radar backscatter signals, with the aim of classifying breast tumours in terms of size and type (malignant or benign), and ultimately attempting to determine the stage of development based on shape (spiculated, microlobulated, macrolobulated and smooth) in homogeneous breast tissue. A set of 352 tumour models of different sizes and shapes was classified using different combinations of three feature extraction methods (PCA, ICA and DWT) and three classifiers (LDA, QDA and SVM).

As far as PCA is concerned, it was found that 30 principal components offer a good compromise between classification accuracy and computational time. For DWT the best performance was obtained when extracting the features from the lowest frequency subband after applying DWT twice with a *Coiflet 5* wavelet.

Comparing the classification approaches, the SVM outperforms both LDA and QDA for all the binary-based architectures and for the comparable architectures DFS and DFSFS (used with LDA and QDA exclusively). SVM is the best classifier in this study because it is able to handle nonlinear relations between measurements and their classes by mapping the measurements in a higher dimension feature space,

allowing for an optimised classification, whereas the dimension feature space of LDA and QDA is limited by the dimensions of the measurements.

Regarding feature extraction methods, both PCA and DWT exhibited better performance of the classifiers in the different classification architectures, with PCA having a slight advantage over DWT. The third feature method extraction, ICA, proved to be the least robust, particularly for the more complex classification architectures.

Overall, the best coarse-shape classifier is obtained with the CSCS architecture, in which DWT and SVM are used, achieving an accuracy as high as 91.19%. The best fine-shape classifier is obtained with the FSFS architecture, in which PCA and SVM are used, achieving an accuracy as high as 75.28%.

With reference to previous results in the literature, specifically in Davis *et al.* [44], the shape classification using PCA and LDA, was accomplished with an accuracy of 86.1% after a previous size classification. In this context the present coarse shape classification of 91.19% obtained with DWT and SVM presents a promising improvement for the classification algorithm described in this chapter.

These results are very promising for improved diagnosis and treatment of early-stage breast cancer within the context of UWB Radar Imaging, under the conditions of the experiment. The next chapter will look at classification of breast tumours in a more realistic scenario as the breast tissue will be modelled with a portion of heterogeneous tissue based on realistic MRI-based models of the breast [83].

Chapter 6

Classification of Breast Tumours in Heterogeneous Breast Tissue

6.1. Motivation and Contributions

In this chapter, the effect of heterogeneity on the classification algorithms presented in the previous chapter is addressed by introducing an amount of fibroglandular breast tissue extracted from 3D MRI models of the breast, taken from the UWCEM Numerical Breast Phantom Repository [83]. This chapter is closely related to the previous chapter as a 3D database of tumours of different sizes and shapes based on GRSs are again modelled with dielectric properties as reported by Lazebnik *et al.* [26-27]. Nonetheless, the inclusion of fibroglandular tissue in the breast models in this chapter presents a much more challenging scenario for classification analysis than the homogeneous breast models used in the previous chapter.

The contributions of this chapter are as follows:

- Use of a larger database with 480 tumour models based on GRSs comprising 30 models for each of the 4 sizes and 4 shapes, with tumours embedded in two different breast model scenarios: (i) breast model with a cluster of fibroglandular breast tissue in a fixed location independent of the tumour location, and, finally, (ii) breast model with a cluster of fibroglandular breast tissue, in a fixed location, possibly overlapping with tumour. The cluster of fibroglandular tissue is extracted from a MRI model of the breast with a high content of adipose tissue compared to fibroglandular tissue.
- Evaluation of three feature extraction methods, PCA, ICA and DWT, with three classification methods, LDA, QDA and SVM, combined with the classification architectures which performed best in the study reported in Chapter 5: CSCS, CSFS, FSCS and FSFS;

- Study of a more challenging heterogeneous scenario in which the one or two clusters of fibroglandular tissue are modelled across a range of different locations within the breast.

The results of this work are presented in Conceição *et al.* [53] and O'Halloran *et al.* [56].

6.2. Tissue Modelling

The numerical simulation for this chapter is similar to that described in the previous chapter, in Section 5.2. For the sake of brevity only the differences between the two simulations are described in this section.

The breast models in which simulations are carried out are based on the models presented in the precedent chapter, however one or two fibroglandular clusters are introduced in different locations. The inclusion of clusters of fibroglandular breast tissue in the breast models allow for the analysis of the effect of dielectric heterogeneity on the tumour classification methods. This is an important consideration as the inclusion of heterogeneity permits a more realistic test platform for the classification algorithms. The simulations are carried out as follows:

- The Total Field (TF) is located at the centre of the Scattered Field (SF) and is represented by a larger cube with 50 mm on each side. Figure 6.1, adapted from [50-51], shows a representation of the TF/SF grid, with the location of the origin of the first incident plane wave and observer point (filled circle) as well as the position of the tumour. All four observation points are represented by small circles.
- A cluster of fibroglandular breast tissue is a block volume of 1 cm^3 representing a well-defined fibroglandular structure taken from a mostly fatty breast model [83]. This cluster is located within the TF region in different locations, and quantities, depending on the breast model.

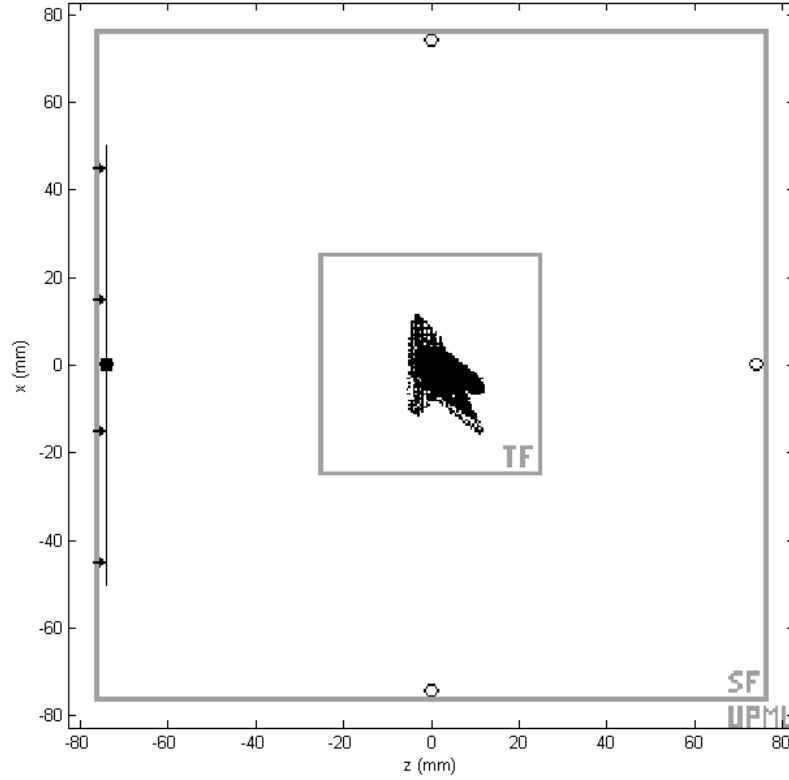


Figure 6.1. Cross-section of the 3D FDTD space lattice partitioned into Total Field (TF), Scattered Field (SF) and UPML regions, for a heterogeneous breast model. In this example, the target, a spiculated tumour located at the centre of the TF, is illuminated by a pulsed plane wave propagating in the $+z$ direction (represented by a dark line) and backscatter is recorded at the first observer location: $(0,0,-74)$ mm (represented by a filled circle). All four observation points are represented by small circles in the image.

The block of fibroglandular breast tissue is extracted from a geometrically and dielectrically accurate 3D breast model provided by the UWCEM Numerical Breast Phantom Repository at the University of Wisconsin [83], which provides the spatial distribution of the different constituent tissues within the breast. These are 3D MRI-derived models taken from patients lying in the prone position. As before, the different tissues within the breast are mapped to the dielectric properties from Lazebnik *et al.* [26-27]. For this work, a 1 cm^3 grid is extracted from a mostly fatty breast phantom available in [83], in particular the phantom with ID 071904. The reason why a mostly fatty breast phantom is used for selection of a fibroglandular volume of breast tissue is due to the fact that fibroglandular structures can be isolated more clearly in such a phantom.

The different tissues within the breast models are modelled using Debye parameters for malignant tissue and for homogeneous lossy adipose tissue, as established by Lazebnik *et al.* [26-27]. For the fibroglandular clusters which account for breast heterogeneity, there are three different levels of fibroconnective/glandular

tissue, as established in [83]. The dielectric properties are shown in Table 6.1 below:

Table 6.1. Debye parameters for the FDTD model. Parameters are established in: (1) Lazebnik et al. [26-27] and (2) UWCEM website [83].

Tissue	ϵ_{∞}	$\Delta\epsilon$	$\sigma_s \text{ (sm}^{-1}\text{)}$	$\tau \text{ (ps)}$
Lossy adipose tissue ⁽¹⁾	3.140	1.708	0.036	14.65
Fibroconnective/glandular – low ⁽²⁾	9.941	26.60	0.462	10.90
Fibroconnective/glandular – median ⁽²⁾	7.821	41.48	0.713	10.66
Fibroconnective/glandular – high ⁽²⁾	6.151	48.26	0.809	10.26
Malignant tissue ⁽¹⁾	6.749	50.09	0.794	10.50

Two different configurations of fibroglandular tissue, with increasing levels of difficulty, are considered in this chapter. The first scenario, represented by Models I and II, have a cluster of fibroglandular tissue which is kept located in one place within the TF region of the breast model. For the second scenario, the best feature extraction and classifier methods are picked from the results from the first scenario and are applied to Models III and IV. For Models III and IV, there are, respectively, one and two clusters of fibroglandular tissue which can be randomly and independently located within the TF region of the breast model. Further detail on the two scenarios is detailed in the following text.

6.2.1. Modelling with Fixed Location of Fibroglandular Tissue

For this part of the results, the tumour database is increased from 352 to 480 models, so that there are 30 tumour models for each of the four sizes and four shapes of tumours considered in Chapter 3. There are two different Models (I and II), represented through a FDTD simulation.

For Model I the portion of fibroglandular breast tissue is located within the cubic TF region in one of its vertices. For Model II the same block of fibroglandular breast tissue is also located within the TF region, at a distance of $\sqrt[3]{5} \text{ mm}$ from one vertex of the cubic TF region, i.e. the block is moved 5 mm in each of the X, Y and Z axes towards the centre of the TF region. In Figure 6.2, a representation of a sample of benign tumours in Model I and Model II is shown.

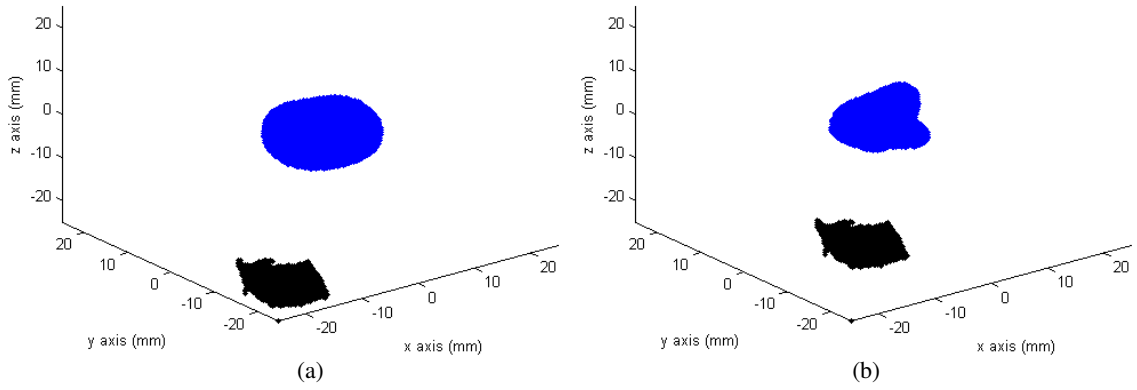


Figure 6.2. Samples of different Gaussian Random Spheres representing benign tumours. The (a) smooth and (b) macrolobulated tumour models are represented in Model I and Model II, respectively. Tumours are represented in blue and fibroglandular clusters in black.

In the first modelling scenario, the feature extraction methods (PCA, ICA and DWT) and the classification approaches (LDA, QDA and SVM) are applied to the backscattered signals of the tumour models in the same fashion as in the previous chapter. However, only the following classification architectures are utilised, as these were the ones found to perform the best and most robustly in Chapter 5: CSCS, CSFS, FSCS and FSFS.

6.2.2. Modelling with Varying Location of Fibroglandular Tissue

For the second modelling scenario, the tumour database includes 160 models, 10 tumour models for each of the four sizes and four shapes considered in Chapter 3. Two different Models (III and IV) are considered for this section of results.

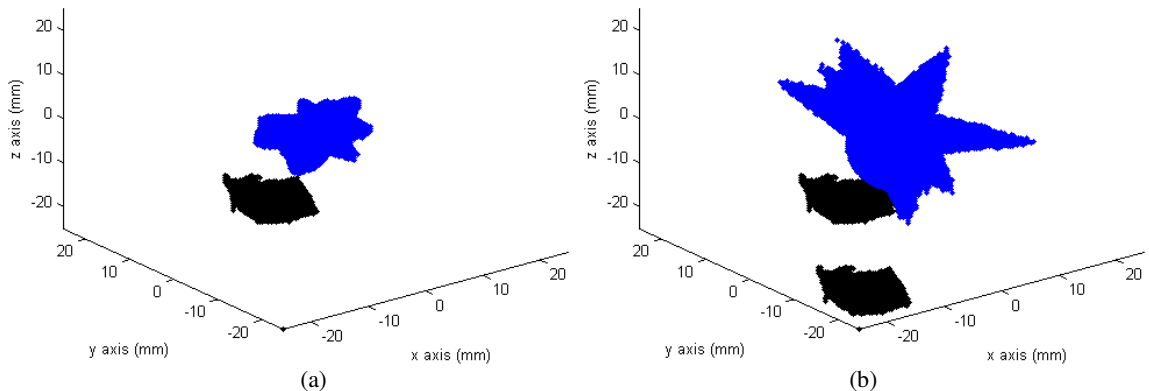


Figure 6.3. Samples of different Gaussian Random Spheres representing malignant tumours. The (a) microlobulated and (b) spiculated (10 spicules) tumour models are represented in samples of Model III and Model IV, respectively. Tumours are represented in blue and fibroglandular clusters in black.

For Model III the block of fibroglandular breast tissue is randomly located in one out of ten locations spread within the TF region. For Model IV the TF region is modelled with two independent blocks of fibroglandular breast tissue which can be located in one out of ten random pairs of locations. In Figure 6.3, a representation of a sample of malignant tumours in Model III and Model IV is shown.

For Models III and IV, and since PCA provides the best results when combined with SVM regarding the first modelling scenario, only the results for this combination will be reported in the interest of brevity. For the second set of experiments reported on here, the same four multi-stage classification architectures (CSCS, CSFS, FSCS and FSFS) are used.

It should be noted that for the simulations in both Chapters 5 and 6, the tumour response is not isolated from the rest of the signal. As an example, Figure 6.4 presents the backscattered signals from a smooth tumour with a radius of 2.5 mm embedded in different breast models: a homogeneous breast model as described in Chapter 5, and in heterogeneous Model II and Model IV as described in this chapter.

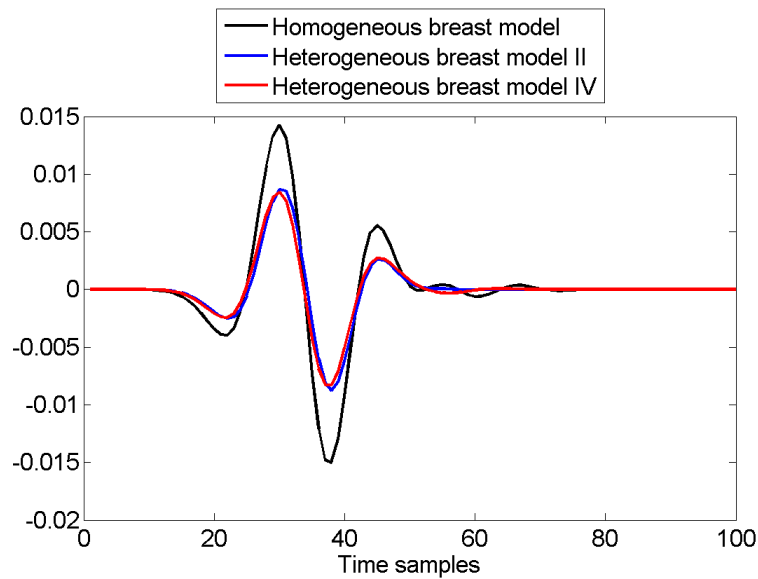


Figure 6.4. Sample of backscattered signals for smooth tumour with a 2.5 mm radius in a homogeneous breast model (black), in a heterogeneous breast model with fixed location of fibroglandular cluster (Model II in blue) and in a heterogeneous breast model with varying location of two fibroglandular cluster (Model IV in red).

6.3. Results and Discussion

6.3.1. Classification with Fixed Location of Fibroglandular Tissue

The performance results for the simulations for Models I and II are shown in Table 6.2 and Table 6.3, respectively. In these tables, the classification is ordered by the four classification architectures composed of partial size and shape classifiers. As before, partial size and shape classifiers provide classification of tumours in terms of either size or shape only, respectively, while the size-then-shape classifiers attempt to classify a tumour in both size and shape. Finally, the best performances for each stage of the classification architectures are highlighted.

Table 6.2. Classification performance for Model I.

Architectures of classifiers	Feature extraction methods	Partial size classifier (%)			Partial shape classifier (%)			Size-then-shape classifier (%)		
		LDA	QDA	SVM	LDA	QDA	SVM	LDA	QDA	SVM
Coarse-Size-Coarse-Shape	PCA	93.75	91.46	94.79	84.79	83.33	92.71	79.49	76.21	87.88
	ICA	92.29	90.62	93.12	79.17	77.50	91.46	73.06	70.23	85.17
	DWT	93.33	91.87	94.37	83.96	82.50	91.46	78.36	75.80	86.31
Coarse-Size-Fine-Shape	PCA	93.75	91.46	94.79	61.25	54.37	74.58	57.42	49.73	70.70
	ICA	92.29	90.62	93.12	54.17	46.67	72.08	49.99	42.29	67.13
	DWT	93.33	91.87	94.37	58.96	55.83	74.58	55.03	51.30	70.39
Fine-Size-Coarse-Shape	PCA	83.12	73.33	85.83	86.25	82.08	92.71	71.69	60.19	79.57
	ICA	79.79	73.12	81.46	82.71	76.04	90.62	65.99	55.60	73.82
	DWT	84.17	75.00	85.21	86.87	81.25	92.71	73.12	60.94	78.99
Fine-Size-Fine-Shape	PCA	83.12	73.33	85.83	68.12	57.71	73.96	56.63	42.32	63.48
	ICA	79.79	73.12	81.46	58.75	49.37	71.87	46.88	36.10	58.55
	DWT	84.17	75.00	85.21	66.67	56.87	75.83	56.11	42.66	64.62

A first analysis of the two tables obtained in this chapter, compared with Table 5.3 in Chapter 5, indicates that the introduction of a portion of fibroglandular breast tissue in a fixed location within the breast model does not influence greatly the classification results, suggesting that the system is quite robust under the conditions in this study. A more detailed analysis of the two tables is presented in the following text.

Table 6.3. Classification performance for Model II.

Architectures of classifiers	Feature extraction methods	Partial size classifier (%)			Partial shape classifier (%)			Size-then-shape classifier (%)		
		LDA	QDA	SVM	LDA	QDA	SVM	LDA	QDA	SVM
Coarse-Size-Coarse-Shape	PCA	92.71	91.87	94.79	83.96	82.71	92.08	77.84	75.99	87.29
	ICA	92.29	90.83	94.17	78.33	77.50	91.46	72.29	70.39	86.12
	DWT	93.12	90.62	94.79	85.42	82.50	91.04	79.54	74.76	86.30
Coarse-Size-Fine-Shape	PCA	92.71	91.87	94.79	59.79	55.00	75.42	55.43	50.53	71.49
	ICA	92.29	90.83	94.17	52.92	47.08	74.37	48.84	42.77	70.04
	DWT	93.12	90.62	94.79	59.79	54.37	74.17	55.68	49.28	70.30
Fine-Size-Coarse-Shape	PCA	82.71	75.21	85.62	86.25	83.33	91.25	71.33	62.67	78.13
	ICA	79.37	72.92	81.46	82.08	77.50	88.96	65.15	56.51	72.46
	DWT	83.12	73.12	85.83	86.87	81.87	92.50	72.21	59.87	79.39
Fine-Size-Fine-Shape	PCA	82.71	75.21	85.62	66.67	60.00	75.42	55.14	45.12	64.57
	ICA	79.37	72.92	81.46	58.33	50.62	71.04	46.30	36.91	57.87
	DWT	83.12	73.12	85.83	66.04	57.92	75.42	54.90	42.35	64.73

Regarding Table 6.2, for both partial size and shape classifiers, the best results are obtained with the SVM classifier. On average, SVM outperforms LDA and QDA by approximately 5.8% and 11.2%, respectively. In terms of the feature extraction methods, PCA and DWT allow for very similar classification results, with PCA presenting a small advantage. On average, PCA allows for a classification that outperforms ICA and DWT by approximately 3.5% and less than 0.1%, respectively. It is also worth mentioning that the best coarse-shape and fine-shape classifications are achieved with both CSCS and FSCS, and FSFS, respectively.

Similarly, the results presented in Table 6.3 show that the SVM method allows for best partial size and shape classifications, where the fibroglandular tissue is located closer to the tumour. On average, SVM outperforms LDA and QDA by approximately 6.35% and 11.1%, respectively. In terms of the feature extraction method, PCA and DWT also allow for very similar classification results as for Table 6.2, with PCA outperforming DWT. On average, PCA allows for a classification that outperforms ICA and DWT by approximately 3.6% and 0.4%, respectively. It is also worth mentioning that the best coarse-shape and fine-shape classifications are achieved with FSCS and both CSFS and FSFS, respectively.

6.3.2. Classification with Varying Location of Fibroglandular Tissue

In the breast models used for the work reported in this section, the fibroglandular

tissue does not have a fixed position within the homogenous breast model. For Model III there is one cluster of fibroglandular tissue which is randomly located in one of ten different locations within the breast whereas for Model IV there are two independent clusters of fibroglandular tissue which are randomly located in one of ten different pairs of locations. The performance results of the classifiers for Models III and IV are shown in Table 6.4, in the same fashion as for the previous section.

As noted earlier, in this section only results for SVM combined with PCA are reported, for the sake of brevity. This choice is motivated by the results from the previous section, where SVM was found to be the best classifier and PCA outperformed the other feature extraction methods.

Table 6.4. Classification performance for Models III and IV (one and two randomly-located heterogeneous clusters, respectively).

Architectures of classifiers	Number of heterogeneous clusters (Model)	Partial size classifier (%)	Partial shape classifier (%)	Size-then-shape classifier (%)
Coarse-Size-Coarse-Shape	one (III)	96.25	91.87	88.43
	two (IV)	91.87	87.50	80.39
Coarse-Size-Fine-Shape	one (III)	96.25	71.87	69.18
	two (IV)	91.87	65.00	59.72
Fine-Size-Coarse-Shape	one (III)	83.12	90.62	75.33
	two (IV)	80.00	85.00	68.00
Fine-Size-Fine-Shape	one (III)	83.12	68.12	56.63
	two (IV)	80.00	61.25	49.00

Table 6.4 indicates that a small decrease is observed when classifying tumours embedded in breast models with varying locations, when one heterogeneous cluster is included (Model III), compared to results in the previous section (Models I and II).

It is also observed that there is further performance decrease when the number of fibroglandular clusters increases to two, in Model IV. This is particularly noticeable in the fine shape classifiers where the difference can be as high as 6.87% (71.87% versus 65.00% for CSFS and 68.12% versus 61.25% for FSFS).

6.4. Conclusions

A database of different tumours with varying size and shape are classified in two different scenarios, with variations regarding breast heterogeneity, using the same

feature extraction and classification algorithms as in Chapter 5. Four breast models were considered: Model I includes a single cluster of fibroglandular tissue which is isolated from the tumour, Model II has a single cluster of fibroglandular tissue which may overlap with the tumour, Model III has a single cluster of fibroglandular tissue that varies in its location within the breast and Model IV includes two clusters of fibroglandular tissue in varying locations within the breast.

Three different feature extraction methods (PCA, ICA and DWT) were used to extract the most significant features, which in turn were used with three classifiers (LDA, QDA and SVM), and using the four classification architectures that provided the best results from studies using homogeneous breast tissue models in Chapter 5 (CSCS, CSFS, FSCS and FSFS).

For the first scenario of breast models with fibroglandular tissue fixed in one location, and for both partial size and shape classifiers, SVM was the most accurate classifier, being the more robust to breast model changes. For Model I, the coarse and fine shape classification achieved a performance as high as 92.71% and 75.83%, respectively. For Model II the coarse and fine shape classification achieved a performance as high as 92.50% and 75.42%, respectively. For both Models I and II, PCA allows for the best classifications (closely followed by DWT), while ICA provides poorer classification performance.

For Model III, where the fibroglandular tissue randomly assumed one of 10 different locations within the homogeneous breast model, the coarse and fine shape classification achieved a performance as high as 91.87% and 71.87%, respectively. Finally, for Model IV (two fibroglandular clusters randomly located within the breast), the coarse and fine shape classification achieved a performance as high as 87.50% and 65.00%, respectively.

Overall, the best classification accuracy results are comparable, independent of the breast model in which the tumours are embedded. In fact, the introduction of a fixed structure of fibroglandular tissue does not introduce any significant changes in the classification performance when compared to results for a homogeneous breast model. However, the introduction of cluster(s) of fibroglandular tissue in varying locations within the breast induces a decrease in the system performance, particularly for the fine shape classifier, more noticeable when two varying fibroglandular clusters are introduced in the breast model at once.

Comparing to previous results in the literature, specifically in Chen *et al.* [45-46]

a coarse shape classification was accomplished with an accuracy of 86.7% using a fixed heterogeneous 2D breast model. Comparable results in Chapter 6 have indicated a coarse shape classification between 87.50% and 92.71% between the models with fixed and random location of fibroglandular cluster(s). Even though the breast models present more heterogeneity in [45-46], it must also be noted that the breast models in Chapter 6 are considered in 3D as well as tumours of four different sizes (versus tumours with a constant radius size in [45-46]).

The results presented in this chapter are promising for breast tumour classification, within the context of UWB radar imaging, since the feature extraction methods and the classification methods analysed in the previous chapter have shown to behave quite robustly with the tumour databases of different sizes, and with the inclusion of clusters of fibroglandular tissue in fixed locations or in a more realistic scenario in which the clusters can take varying locations within the breast. Future work may include the inclusion of multiple clusters of fibroglandular tissue within the breast model for even more realistic simulations.

Chapter 7

Conclusions and Discussion

7.1. Summary of Work

In this thesis, a number of aspects of Ultra Wideband Radar system design for breast imaging were addressed, with a view to the development of systems for robust detection and classification of breast cancer.

The physiological background regarding the breast and tumour development was first presented. Historical research related to the measurement of dielectric properties of the human female breast, as well as breast tumours was reviewed. In particular, it was reported that recent findings from Lazebnik *et al.* [26-27] have identified overlapping dielectric properties between normal fibroglandular breast tissues and tumour tissues (both benign and malignant), which renders the task of tumour detection based on dielectric properties more difficult, and which prompted much of the work developed in this thesis. The development of UWB Radar algorithms for early-stage breast cancer detection was also reviewed, as well as the state of the art in UWB-based classification algorithms for malignant and benign tumours, in order to set the context for the research described in this thesis.

One of the largest current databases of tumours for tumour classification using UWB was created for this work, comprising 480 tumour models based on Gaussian Random Spheres ranging between four sizes and four shapes, which included two types of benign and two types of malignant tumours. This database was used as the simulation platform for subsequent tests in the thesis.

Two different approaches for the antenna configuration in a UWB imaging system, planar and circular, were compared and optimised in different breast model scenarios, with the aid of a number of performance metrics. Firstly, both antenna configurations were optimised for tumours of varying sizes, between 2 and 10 mm, and secondly, the antenna configurations were optimised for tumours varying between three different locations within the breast. Both antenna configurations were

also tested in terms of their robustness to increases in the variability of the dielectric constants of the tissue models.

Finally, classification of tumours was addressed by the use of a number of different feature extraction and classification algorithms, which operate on the Radar Target Signatures (RTS) produced by tumours. Three feature extraction methods were investigated: Principal Component Analysis (PCA), Independent Component Analysis (ICA) and Discrete Wavelet Transform (DWT). Three classifiers were also considered: Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA) and Support Vector Machines (SVM). The feature extraction and classification algorithms were evaluated in various combinations in order to determine the best approach for tumour classification. The use of different multi-stage classification architectures was also proposed to further optimise classification performance. Classification was obtained for both homogenous and heterogeneous breast models; in particular, the breast models that include heterogeneity were tested in two different scenarios, one with a single fibroglandular tissue cluster in a fixed location, and a second scenario with fibroglandular tissue cluster(s) with varying locations within the breast model. The inclusion of fibroglandular tissue in the breast models presents a more challenging test environment than has generally been considered in the existing literature.

7.2. Conclusions and Discussion

This section discusses the primary results arising from the work presented in this thesis, organised according to the main individual components of research.

7.2.1. Modelling of Tumours

An algorithm for creation of tumour models based on Gaussian Random Spheres (GRSs) was adapted and developed, and different types of benign and malignant tumours of different sizes were created to complete a large database that served as the test set for classification algorithms, also addressed in this thesis. This is one of the largest databases of breast tumour models based on GRSs that has been reported so far, comprising 480 models of four different sizes and four different shapes, in which two shapes represent benign tumours and the other two represent malignant

tumours. Furthermore, the range of tumour shapes contained in the database was extended to allow for differentiation of benign tumours into smooth and macrolobulated categories; this is an important addition, since the macrolobulated tumour may be an indicator of pre-malignancy in a clinical setting.

7.2.2. Antenna Configurations

Planar and circular antenna configurations were compared in detail for tumours of varying size, and varying location in the breast, and the optimum number of antennas was found for each condition. The antenna configurations, were tested using a number of performance metrics: Signal to Clutter Ratio within breast (SCRwb), Signal to Clutter Ratio between breast (SCRbb), Signal to Mean Ratio, (SMR), all of which reflect the strength of the tumour energy response against clutter, and Full Width Half Maximum (FWHM) extent of the tumour response, and the distance between the actual location of the tumour and the location of its peak location (M_{diff}), all of which reflect how correctly the location of the tumour is determined.

Simple tumours as small as 2 mm were detected successfully in the first described scenario. Comparing the two antenna configurations, it was found that the circular configuration allowed for better performance in detecting small tumours; for example, for a 6 mm tumour, the SCRwb for the circular configuration was up to 4 dB better than the SCRwb for the planar configuration. Furthermore, the circular configuration was more robust to changes in the dielectric variability. For the circular configuration, the centre of the tumour in the UWB image is closer to its correct position compared to the planar configuration, for which the tumour appears to be closer to the skin layer than it actually is. This indicates that having antennas arranged around the tumour allows for better detection of its location, compared to having antennas on only one side of the tumour.

For the second set of experiments with tumours located in different parts within the breast, the tumour again appears closer to the skin than it actually is in the planar configuration, as found in the first set of experiments with varying tumour sizes. However, it was observed that clutter is less significant in the planar images than in the circular images. The planar configuration shows better performance in terms of both SCRwb and SMR metrics, except for the case where the tumour is located

furthest from the centre of the breast, due to increased average distance between the antennas and the tumour location. However, the circular configuration allows for better localization of the tumour, indicated by better M_{diff} and FWHM results for this configuration. For the circular configuration, the performance metrics remained relatively constant independent of the tumour location. In general, it was also found that both planar and circular antennas were quite robust to dielectric variance changes.

The optimum number of antennas was found to 16 antennas for tumours of different sizes, and 17 antennas for tumours of different locations, given the results of the performance metrics and visual inspection. Therefore, on the order of 17 antennas should be a useful guide figure for the design of UWB imaging systems. Although the circular configuration was found to be the configuration with better performance, it should also be noted that this configuration should not be used exclusively, as it does not allow for the effective examination of tumours developing in the underarm region, where a sizeable proportion of breast tumours may occur.

7.2.3. Classification of Tumours in a Homogeneous Tissue

For a database of 352 tumours embedded in a homogeneous breast tissue model, three feature extraction methods (PCA, ICA and DWT) and three classifiers (LDA, QDA and SVM) were investigated. Furthermore, eight different multi-stage classification architectures containing different combinations of size and/or shape classifiers were proposed, for different applications.

Firstly, the LDA and QDA classifiers were compared with the three feature extraction methods (PCA, ICA and DWT). It was found that shape classifiers performed better with a preceding size classifier to “filter” the results. In addition, it was found that binary-based classification architectures performed more robustly. The accuracy of coarse and fine shape classifications were as high as 86.08% and 68.18%, respectively, for PCA and DWT combined with LDA, when a previous fine size classifier was used (corresponding to the FSCS and FSFS multi-stage classification architectures presented in Chapter 5). When considering the overall accuracy of a size-then-shape cascade classifier, the accuracy of the cascade classifiers worsens with increasing number of stages; this may not be unexpected,

since fine classification is a more challenging problem than coarse classification. It was observed that, overall, LDA outperforms QDA as a classification algorithm.

Classification based on the use of Support Vector Machine (SVM) was also examined for the same feature extraction methods. In this case, the subset of multi-stage classification architectures that performed the best in the initial set of experiments was used. Again, it was found that both PCA and DWT result in higher coarse and fine shape performance, 91.19% and 75.28%, when combined with architectures with previous size classifications, CSCS and FSFS respectively. SVM was also found to significantly outperform the LDA and QDA approaches. PCA and DWT provided similar classification performance, while ICA was significantly poorer in extracting relevant features. While classification errors at any particular stage in the multi-stage architectures did not show any particular bias (i.e. the number of misclassification for one class was generally similar to the errors for the other classes), there were some exceptions. For example, the similarity between some smooth and macrolobulated tumours resulted in greater numbers of classification errors in these cases.

In terms of the best combination of feature extraction method and classifier for size classification for the homogeneous tissue case, both coarse and fine classification performed best for PCA combined with SVM, achieving a performance of 94.89% and 86.93%, respectively. For shape classification, the best combination was found to be DWT combined with SVM, which provided classification performance of 91.19%. If extra granularity in shape classification is required, the best configuration is PCA combined with SVM, which yielded an overall performance of 75.28%.

7.2.4. Classification of Tumours in Heterogeneous Tissue

Research into the classification of tumours in homogeneous tissue summarised in the previous section was extended to the scenario where one or more clusters of fibroglandular tissue was included in the breast. The robustness of the classification algorithms was tested using a larger tumour database comprising 480 models, in more realistic breast models which accounted for dielectric heterogeneity. Two scenarios are tested; firstly, one in which a single fibroglandular cluster is located in

a fixed placed within the homogeneous breast model (both independent of, and overlapping with the tumour), and a second scenario in which one or two fibroglandular clusters vary location randomly within the breast.

Classification performance remained high for the breast models with fixed fibroglandular clusters (when compared to performance with homogeneous breast models), showing very small average variations between the two heterogeneous breast scenarios. The best coarse classification of tumours varied between 92.71% (in CSCS and FSCS) and 92.50% (in FSCS), while fine classification of tumours varied between 75.83% (in FSFS) and 75.42% (in CSFS and FSFS). As before, fine size or shape classifications result in lower performance than coarse classifications, because of the more challenging classification problem. The best performance was obtained using SVM and PCA or DWT as the classifier and feature extraction methods.

The introduction of one or two randomly-located fibroglandular clusters within the model of homogeneous breast tissue had a slightly larger effect on the robustness of the algorithms. PCA combined with SVM were used for classification, as these algorithms were generally found to be the best performing from the previous experiments. The best coarse shape classification of tumours was 91.87% and 87.50%, when one or two clusters of fibroglandular tissue were present in the breast model, respectively. This was obtained using the CSCS architecture. Furthermore, the best fine shape classification of tumours was 71.87% and 65.00% (in CSFS), when one or two clusters of fibroglandular tissue were modelled, respectively.

The development of classification algorithms for heterogeneous breast models addressed in this thesis presents a much more challenging problem than the homogeneous breast tissue case. This is particularly significant in light of recent findings suggesting that there is significant overlap between the dielectric properties of normal fibroglandular tissue, and tumour tissue. The results obtained are promising, and point to the basic viability of classification of tumours based on radar target signatures obtained using a UWB imaging system.

7.2.5. Summary of contributions

The main contributions of the thesis are briefly summarised in the following bullet points:

- Creation of a large tumour database of 480 tumours comprising four sizes and four shapes of tumour models based on GRSs; this is one of the largest databases based on GRSs reported thus far.
- Inclusion of macrolobulated as an additional type of benign tumour, to model the case of a tumour in the pre-malignant stage.
- Evaluation of circular and planar antenna configuration for a number of different scenarios, and establishment of the circular configuration as the one which generally performed better.
- Comprehensive evaluation of a range of feature extraction and classification algorithms for tumour categorisation, and proposal of a range of multi-stage architectures. The best performance obtained was 91.19% using DWT as the feature extraction method and SVM as the classifier, for coarse classification.
- Evaluation of classifiers in the much more challenging scenario of heterogeneous breast models, containing one or two clusters of fibroglandular tissue in fixed and varying locations. Coarse shape classification performance as high as 92.71% was obtained, with fine shape classification performance as high as 75.83%, for fixed locations of fibroglandular clusters. For varying locations of fibroglandular clusters, coarse shape classifier performed as high as 91.87% and 87.50%, and the fine shape classifier performed as high as 71.87% and 65.00%, with one and two clusters of fibroglandular tissue, respectively, using PCA and SVM.

7.3. Suggestions for Future Work

The work developed in this thesis presents many areas which can be further investigated in order to improve the detection and classification of early breast tumours. Suggestions for future work as a follow-on to the work presented in this thesis include the following topics.

The present work included a number of realistic modelling scenarios, particularly those involving fibroglandular tissue in heterogeneous breast models. The range of models used for system development and evaluation could be further

enhanced, for example, even more clusters of fibroglandular tissue, with randomly varying locations, may be included. The range of breast models may be extended to include a full range of tissue topologies, ranging from very fatty breasts to very dense breasts. The modelling of dense breasts is of particular interest for the development of systems for screening women at an earlier age, as the breast tissue is denser in younger women, and the imaging problem is more significant.

A different approach to modelling tumours may also be considered, for example, based on the models presented in Chen *et al.* [45-47, 95, 98]. In their algorithms, the tumours are modelled through a 2D polygonal approximation of tumour boundaries from X-Ray mammograms [41]. This work could be extended to a more realistic 3D environment, and would allow for the creation of a sizeable database for further classification purposes.

Additional features for classification could be developed and evaluated. For example, the DWT feature extraction framework may be extended to include information over a broader part of the Radar Target Signature (RTS) spectrum. The DWT could be replaced by the Wavelet Packet Transform, which provides greater resolution in frequency than the DWT, and opens up the possibility for a more powerful representation of the tumours by selection of the best basis based on e.g. entropy or an alternative criterion. More powerful features may improve the performance of the classification system by providing better discrimination in cases where classification shows scope for improvement, e.g. in the case of distinguishing between smooth and macrolobulated tumours discussed above.

Finally, in addition to the classification algorithms examined in this thesis, other classification algorithms may be considered, such as the following: Multi-Layer Perceptrons, Genetic Algorithms and Spiking Neural Networks. The use of genetic algorithms or Spiking Neural Networks may allow for evolution of classifier parameters based on more “intelligent” signal processing algorithms. It may also permit greater levels of adaptivity to allow classifiers to better adjust to changes in tissue characteristics, as opposed to most of the standard classification algorithms whose performance is largely dependent on the range of data available to train them.

References

1. *Cancer Facts & Figures*, American Cancer Society, Atlanta, 2008.
2. *Europa Donna Ireland - The Irish Breast Cancer Campaign*. 2010. Last accessed: 22/07/2010; Available from: www.europadonnaireland.ie.
3. A. Langtry, *Understanding Cancer of the Breast*, J. Kelly (Editor). Irish Cancer Society. Dublin, Ireland, 2008.
4. T. Gorey, J. Kennedy, F. Flanagan, C. Faul and M. Murray (2006) *The Breast in Health and Illness - An Information Guide for Patients and Their Carers (DVD)*, pp., AstraZeneca Pharmaceuticals (Irl.) Ltd.
5. *The National Cancer Registry Ireland*. Last accessed: 22/07/2010; Available from: www.ncri.ie.
6. R. E. Bird, T. Wallace and B. Yankaskas, "Analysis of Cancers Missed at Screening Mammography", *Radiology*, Vol. 184, pp. 613-617, 1992.
7. P. T. Huynh, A. M. Jarolimek and S. Daye, "The False-Negative Mammogram", *RadioGraphics*, Vol. 18, pp. 1137-1154, 1998.
8. S. L. Nass, I. C. Henderson and J. C. Lashof, *Mammography and Beyond: Developing Technologies for the Early Detection of Breast Cancer*. National Academy Press, 2001.
9. J. G. Elmore, M. B. Barton, V. M. Moceris, S. Polk, P. J. Arena and S. W. Fletcher, "Ten-Year Risk of False Positive Screening Mammograms and Clinical Breast Examinations", *The New England Journal of Medicine*, Vol. 338, No. 16, pp. 1089-1096, 1998.
10. G. Bindu, A. Lonappan, V. Thomas, C. K. Aanandan and K. T. Mathew, "Active Microwave Imaging for Breast Cancer Detection", *Progress In Electromagnetics Research*, Vol. 58, pp. 149-169, 2006.
11. C. K. Kuhl, S. Schrading, C. C. Leutner, N. Morakkabati-Spitz, E. Wardelmann, R. Fimmers, W. Kuhn and H. H. Schild, "Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer", *Journal of Clinical Oncology*, Vol. 23, No. 33, pp. 8469-8476, 2005.

12. S. C. Hagness, A. Taflove and J. E. Bridges, "Two Dimensional FDTD Analysis of a Pulsed Microwave Confocal System for Breast Cancer Detection: Fixed-Focus and Antenna-Array Sensors", *IEEE Transactions on Biomedical Engineering*, Vol. 45, pp. 1470-1479, 1998.
13. J. E. Bridges, S. C. Hagness, A. Taflove and M. Popovic, *Microwave Discrimination between Malignant and Benign Breast Tumours*, Patent No.:US 6,431,550 B1, Assignee: L.L.C. Interstitial, USA, 2002
14. F. MacDonald and C. H. J. Ford, *Molecular Biology of Cancer*. BIOS Scientific Publishers Limited. Oxford, 1997.
15. W. T. Joines, "Frequency-Dependent Absorption of Electromagnetic Energy in Biological Tissue", *IEEE Transactions on Biomedical Engineering*, Vol. BME-31, No. 1, pp. 17-20, 1984.
16. R. Pethig, "Dielectric Properties of Biological Materials: Biophysical and Medical Applications", *IEEE Transactions on Electrical Insulation*, Vol. E1-E19, No. 5, pp. 453-474, 1984.
17. S. S. Chaudhary, R. K. Mishra, A. Swarup and J. M. Thomas, "Dielectric Properties of Normal and Malignant Human Breast Tissues at Radiowave and Microwave Frequencies", *Indian Journal of Biochemistry & Biophysics*, Vol. 21, pp. 76-79, 1984.
18. W. T. Joines, Y. Zhang, C. Li and R. L. Jirtle, "The Measured Electrical Properties of Normal and Malignant Human Tissues from 50 to 900 MHz", *Medical Physics*, Vol. 21, No. 4, 1994.
19. A. J. Surowiec, S. S. Stuchly, J. R. Barr and A. Swarup, "Dielectric Properties of Breast Carcinoma and the Surrounding Tissues", *IEEE Transactions on Biomedical Engineering*, Vol. 35, No. 4, pp. 257-263, 1988.
20. A. M. Campbell and D. V. Land, "Dielectric Properties of Female Human Breast Tissue Measured *In Vitro* at 3.2 GHz", *Physics in Medicine and Biology*, Vol. 37, No. 1, pp. 193-210, 1992.
21. C. Gabriel, S. Gabriel and E. Corthout, "The Dielectric Properties of Biological Tissues: I. Literature Survey", *Physics in Medicine and Biology*, Vol. 41, pp. 2231-2249, 1996.
22. S. Gabriel, R. W. Lau and C. Gabriel, "The Dielectric Properties of Biological Tissues: II. Measurements in the Frequency Range 10 Hz to 20 GHz", *Physics in Medicine and Biology*, Vol. 41, pp. 2251-2269, 1996.

23. S. Gabriel, R. W. Lau and C. Gabriel, "The Dielectric Properties of Biological Tissues: III. Parametric Models for the Dielectric Spectrum of Tissues", *Physics in Medicine and Biology*, Vol. 41, pp. 2271-2293, 1996.
24. J. Jossinet, "The Impedivity of Freshly Excised Human Breast Tissue", *Physiological Measurement*, Vol. 19, No. 1, pp. 61-75, 1998.
25. J. Jossinet and M. Schmitt, "A Review of Parameters for the Bioelectrical Characterization of Breast Tissue", *Annals of the New York Academy of Sciences*, Vol. 873, pp. 30-41, 1999.
26. M. Lazebnik, L. McCartney, D. Popovic, C. B. Watkins, M. J. Lindstrom, J. Harter, S. Sewall, A. Magliocco, J. H. Booske, M. Okoniewski and S. C. Hagness, "A Large-Scale Study of the Ultrawideband Microwave Dielectric Properties of Normal Breast Tissue Obtained from Reduction Surgeries", *Physics in Medicine and Biology*, Vol. 52, pp. 2637–2656, 2007.
27. M. Lazebnik, D. Popovic, L. McCartney, C. B. Watkins, M. J. Lindstrom, J. Harter, S. Sewall, T. Ogilvie, A. Magliocco, T. M. Breslin, W. Temple, D. Mew, J. H. Booske, M. Okoniewski and S. C. Hagness, "A Large-Scale Study of the Ultrawideband Microwave Dielectric Properties of Normal, Benign and Malignant Breast Tissues Obtained from Cancer Surgeries", *Physics in Medicine and Biology*, Vol. 52, pp. 6093–6115, 2007.
28. S. C. Hagness, A. Taflove and J. E. Bridges, "Three-Dimensional FDTD Analysis of a Pulsed Microwave Confocal System for Breast Cancer Detection: Design of an Antenna-Array Element", *IEEE Transactions on Antennas and Propagation*, Vol. 47, No. 5, pp. 783-791, 1999.
29. E. C. Fear, X. Li, S. C. Hagness and M. A. Stuchly, "Confocal Microwave Imaging for Breast Cancer Detection: Localization of Tumors in Three Dimensions", *IEEE Transactions on Biomedical Engineering*, Vol. 49, No. 8, pp. 812-822, 2002.
30. P. M. Meaney, K. D. Paulsen, J. T. Chang, M. W. Fanning and A. Hartov, "Nonactive Antenna Compensation for Fixed-Array Microwave Imaging: Part II—Imaging Results", *IEEE Transactions on Medical Imaging*, Vol. 18, No. 6, pp. 508-518, 1999.
31. P. M. Meaney, M. W. Fanning, D. Li, S. P. Poplack and K. D. Paulsen, "A Clinical Prototype for Active Microwave Imaging of the Breast", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 11, pp.

- 1841-1853, 2000.
32. P. M. Meaney, M. W. Fanning, T. Raynolds, C. J. Fox, Q. Fang, C. A. Kogel, S. P. Poplack and K. D. Paulsen, "Initial Clinical Experience with Microwave Breast Imaging in Women with Normal Mammography", *Academic Radiology*, Vol. 14, No. 2, pp. 207-218, 2007.
33. A. E. Souvorov, A. E. Bulyshev, S. Y. Semenov, R. H. Svenson and G. P. Tatsis, "Two-Dimensional Computer Analysis of a Microwave Flat Antenna Array for Breast Cancer Tomography", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 8, pp. 1413-1415, 2000.
34. A. E. Bulyshev, S. Y. Semenov, A. E. Souvorov, R. H. Svenson, A. G. Nazarov, Y. E. Sizov and G. P. Tatsis, "Computational Modeling of Three-Dimensional Microwave Tomography of Breast Cancer", *IEEE Transactions on Biomedical Engineering*, Vol. 48, No. 9, pp. 1053-1056, 2001.
35. Q. H. Liu, Z. Q. Zhang, T. T. Wang, J. A. Bryan, G. A. Ybarra, L. W. Nolte and W. T. Joines, "Active Microwave Imaging I—2-D Forward and Inverse Scattering Methods", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 50, No. 1, pp. 123-133, 2002.
36. P. Kosmas and C. M. Rappaport, "Time Reversal with the FDTD Method for Microwave Breast Cancer Detection", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 53, No. 7, pp. 2317-2323, 2005.
37. P. Kosmas and C. M. Rappaport, "FDTD-Based Time Reversal for Microwave Breast Cancer Detection – Localization in Three Dimensions", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 54, No. 4, pp. 1921-1927, 2006.
38. P. Kosmas and C. M. Rappaport, "A Matched-Filter FDTD-Based Time Reversal Approach for Microwave Breast Cancer Detection", *IEEE Transactions on Antennas and Propagation*, Vol. 54, No. 4, pp. 1257-1264, 2006.
39. L. Sha, E. Renee and B. Stroy, "A Review of Dielectric Properties of Normal and Malignant Breast Tissue", *Proceedings of the IEEE SoutheastCon 2002*, Columbia, SC, USA, pp. 457-462, 2002.
40. G. Bindu and K. T. Mathew, "Characterization of Benign and Malignant Breast Tissues Using 2-D Microwave Tomographic Imaging", *Microwave and Optical Technology Letters*, Vol. 49, pp. 2341-2345, 2007.

41. R. M. Rangayyan, N. M. El-Faramawy, J. E. L. Desautels and O. A. Alim, "Measures of Acutance and Shape for Classification of Breast Tumors", *IEEE Transactions on Medical Imaging*, Vol. 16, No. 6, pp. 799-810, 1997.
42. T. M. Nguyen and R. M. Rangayyan, "Shape Analysis of Breast Masses in Mammograms Via the Fractal Dimension", *Engineering in Medicine and Biology 27th Annual Conference*, Shangai, China, pp. 3210-3213, 2005.
43. D. Guliato, R. M. Rangayyan, J. D. Carvalho and S. A. Santiago, "Polygonal Modeling of Contours of Breast Tumors with the Preservation of Spicules", *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 1, pp. 14-20, 2008.
44. S. K. Davis, B. D. Van Veen, S. C. Hagness and F. Kelcz, "Breast Tumor Characterization Based on Ultrawideband Microwave Backscatter", *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 1, pp. 237-246, 2008.
45. Y. Chen, I. J. Craddock, P. Kosmas, M. Ghavami and P. Rapajic, "Application of the MIMO Radar Technique for Lesion Classification in UWB Breast Cancer Detection", *17th European Signal Processing Conference (EUSIPCO)*, Glasgow, Scotland, pp. 759-763, 2009.
46. Y. Chen, I. J. Craddock, P. Kosmas, M. Ghavami and P. Rapajic, "Multiple-Input Multiple-Output Radar for Lesion Classification in Ultrawideband Breast Imaging", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 4, No. 1, pp. 187-201, 2010.
47. Y. Chen, I. J. Craddock and P. Kosmas, "Feasibility Study of Lesion Classification via Contrast-Agent-Aided UWB Breast Imaging", *IEEE Transactions on Biomedical Engineering*, Vol. 57, No. 5, pp. 1003-1007, 2010.
48. J. Teo, Y. Chen, C. B. Soh, E. Gunawan, K. S. Low, T. C. Putti and S. Wang, "Breast Lesion Classification Using Ultrawideband Early Time Breast Lesion Response", *IEEE Transactions on Antennas and Propagation*, Vol. 58, No. 8, pp. 2604-2613, 2010.
49. R. C. Conceição, M. O'Halloran, M. Glavin and E. Jones, "Comparison of Planar and Circular Antenna Configurations For Breast Cancer Detection Using Microwave Imaging", *Progress In Electromagnetics Research*, Vol. 99, pp. 1-20, 2009.
50. R. C. Conceição, M. O'Halloran, E. Jones and M. Glavin, "Investigation of Classifiers for Early-Stage Breast Cancer Based on Radar Target Signatures",

- Progress In Electromagnetics Research*, Vol. 105, pp. 295-311, 2010.
51. R. C. Conceição, M. O'Halloran, M. Glavin and E. Jones, "Support Vector Machines for the Classification of Early-Stage Breast Cancer Based on Radar Target Signatures", *Progress In Electromagnetics Research B*, Vol. 23, pp. 311-327, 2010.
 52. R. C. Conceição, M. O'Halloran, M. Glavin and E. Jones, "Evaluation of Features and Classifiers for Classification of Early-Stage Breast Cancer", *Journal of Electromagnetic Waves and Applications*, Vol. 25, No. 1, pp. 1-14, 2011.
 53. R. C. Conceição, M. O'Halloran, M. Glavin and E. Jones, "Effects of Dielectric Heterogeneity in the Performance of Breast Tumour Classifiers", *Progress in Electromagnetic Research M*, Vol. 17, pp. 73-86, 2011.
 54. M. O'Halloran, R. C. Conceição, D. Byrne, M. Glavin and E. Jones, "FDTD Modeling of the Breast: A Review", *Progress In Electromagnetics Research B*, Vol. 18, pp. 1-24, 2009.
 55. B. McGinley, M. O'Halloran, R. C. Conceição, F. Morgan, M. Glavin and E. Jones, "Spiking Neural Networks for Breast Cancer Classification Using Radar Target Signatures", *Progress In Electromagnetics Research C*, Vol. 17, pp. 79-94, 2010.
 56. M. O'Halloran, B. McGinley, R. C. Conceição, F. Morgan, M. Glavin and E. Jones, "Spiking Neural Networks for Breast Cancer Classification in a Dielectrically Heterogeneous Breast", *Progress In Electromagnetics Research*, Vol. 113, pp. 413-428, 2011.
 57. R. C. Conceição, D. Byrne, M. O'Halloran, M. Glavin and E. Jones, "Classification of Suspicious Regions within Ultrawideband Radar Images of the Breast", *16th IET Irish Signals and Systems Conference, ISSC 2008*, Galway, Ireland, Vol. 1, pp. 60-65, 2008.
 58. R. C. Conceição, M. O'Halloran, M. Glavin and E. Jones, "Antenna Configurations for Ultra Wide Band Radar Detection of Breast Cancer", *Proceedings SPIE BIOS West*, San José, CA, USA, Vol. 7169, No. 9, pp. [71691M, 12], 2009.
 59. R. C. Conceição, M. O'Halloran, D. Byrne, E. Jones and M. Glavin, "Tumor Classification Using Radar Target Signatures", *Progress In Electromagnetics Research Symposium*, Cambridge, MA, USA, pp. 346-349, 2010.

60. K. I. Bland, E. M. Copeland III, N. E. Davidson, D. L. Page, A. Recht and M. M. Urist, *The Breast: Comprehensive Management of Benign and Malignant Disorders*, Elsevier. Vol. 1,, 2004.
61. J. W. Choi, J. Cho, Y. Lee, J. Yim, B. Kang, K. K. Oh, W. H. Jung, H. J. Kim, C. Cheon, H. Lee and Y. Kwon, "Microwave Detection of Metastasized Breast Cancer Cells in the Lymph Node; Potential Application for Sentinel Lymphadenectomy", *Breast Cancer Research and Treatment*, Vol. 86, pp. 107-115, 2004.
62. CancerHelp UK. *About Cancer*. 2007. Last accessed: 22/10/2007; Available from: <http://www.cancerhelp.org.uk/help/default.asp?page=85>.
63. F. M.ENZINGER and S. W. WEISS, *Soft Tissue Tumors*. Third Edition. Mosby - Year Book, Inc. Missouri, 1995.
64. I. L. Cameron and T. B. Pool, *The Transformed Cell*. Cell Biology, A Series of Monographs. Academic Press. New York, 1981.
65. A. Malich, B. Scholz, A. Kott, M. Facius and D. R. Fischer, "The Impact of Lesion Vascularisation on Tumours Detection by Electrical Impedance Scanning at 200 Hz", *Biomedical Imaging and Intervention Journal*, Vol. 3, No. 4:e33, 2007.
66. CancerHelp UK. *Breast Cancer Section Overview*. 2007. Last accessed: 21/11/2007; Available from: <http://www.cancerhelp.org.uk/help/default.asp?page=3270>.
67. J. M. Dixon (editor), *ABC of Breast Diseases*. Third Edition. Blackwell Publishing Ltd, 2006.
68. R. J. Halter, T. Zhou, P. M. Meaney, A. Hartov, R. J. Barth Jr, K. M. Rosenkranz, W. A. Wells, C. A. Kogel, A. Borsic, E. J. Rizzo and K. D. Paulsen, "The Correlation of *In Vivo* and *Ex Vivo* Tissue Dielectric Properties to Validate Electromagnetic Breast Imaging: Initial Clinical Experience", *Physiological Measurement*, Vol. 30, No. 6, pp. S121-S136, 2009.
69. D. Haemmerich, R. Ozkan, S. Tungjitkusolmun, J. Z. Tsai, D. M. Mahvi, S. T. Staelin and J. G. Webster, "Changes in Electrical Resistivity of Swine Liver after Occlusion and Postmortem", *Medical & Biological Engineering & Computing*, Vol. 40, No. 1, pp. 29-33, 2002.
70. X. Li and S. C. Hagness, "A Confocal Microwave Imaging Algorithm for Breast Cancer Detection", *IEEE Microwave and Wireless Components Letters*,

- Vol. 11, No. 3, pp. 130-132, 2001.
71. E. J. Bond, X. Li, S. C. Hagness and B. D. Van Veen, "Microwave Imaging via Space-Time Beamforming for Early Detection of Breast Cancer", *IEEE Transactions on Antennas and Propagation*, Vol. 51, No. 8, pp. 1690-1705, 2003.
 72. R. Nilavalan, A. Gbedemah, I. J. Craddock, X. Li and S. C. Hagness, "Numerical Investigation of Breast Tumour Detection Using Multi-Static Radar", *IET Electronics Letters*, Vol. 39, No. 25, pp. 1787-1789, 2003.
 73. M. Klemm, I. J. Craddock, J. Leendertz, A.W. Preece and R. Benjamin, "Breast Cancer Detection Using Symmetrical Antenna Array", *Antennas and Propagation, 2007. EuCAP 2007. The Second European Conference*, Edinburgh, UK, pp. 1-5, 2007.
 74. K. S. Yee, "Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media", *IEEE Transactions on Antennas and Propagation*, Vol. 14, No. 3, pp. 302-307, 1966.
 75. A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method*. Second Edition. Artech House. Boston, 2000.
 76. R. Luebbers, F. P. Hunsberger, K. S. Kunz, R. B. Standler and M. Schneider, "A Frequency-Dependent Finite-Difference Time-Domain Formulation for Dispersive Materials", *IEEE Transactions on Electromagnetic Compability*, Vol. 32, No. 3, pp. 222-227, 1990.
 77. E. C. Fear and M. A. Stuchly, "Microwave Detection of Breast Cancer", *IEEE Transactions on Microwave Theory and Techniques*, Vol. 48, No. 11, pp. 1854-1863, 2000.
 78. H. B. Lim, N. T. T. Nhung, E. Li and N. D. Thang, "Confocal Microwave Imaging for Breast Cancer Detection: Delay-Multiply-And-Sum Image Reconstruction Algorithm", *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 6, pp. 1697-1704, 2008.
 79. M. O'Halloran, M. Glavin and E. Jones, "Channel-Ranked Beamformer for the Early Detection of Breast Cancer", *Progress In Electromagnetics Research*, Vol. 103, pp. 153-168, 2010.
 80. S. K. Davis, E. J. Bond, X. Li, S. C. Hagness and B. D. Van Veen, "Microwave Imaging via Space-Time Beamforming for the Early Detection of

- Breast Cancer: Beamformer Design in the Frequency Domain”, *Journal of Electromagnetic Waves and Applications*, Vol. 17, No. 2, pp. 357-381, 2003.
81. M. Klemm, I. J. Craddock, J. A. Leendertz, A. Preece and R. Benjamin, “Improved Delay-and-Sum Beamforming Algorithm for Breast Cancer Detection”, *International Journal of Antennas and Propagation*, Vol. 2008, pp. 1-9, 2008.
82. M. O'Halloran, M. Glavin and E. Jones, “Effects of Fibroglandular Tissue Distribution on Data-Independent Beamforming Algorithms”, *Progress In Electromagnetics Research*, Vol. 97, pp. 141-158, 2009.
83. *University of Wisconsin - Computational Electromagnetics Laboratory (UWCEM)*. Last accessed: 22/09/2010; Available from: <http://uwcem.ece.wisc.edu/>.
84. M. O'Halloran, M. Glavin and E. Jones, “Quasi-Multistatic MIST Beamforming for the Early Detection of Breast Cancer”, *IEEE Transactions on Biomedical Engineering*, Vol. 57, No. 4, pp. 830-840, 2010.
85. M. O'Halloran, M. Glavin and E. Jones, “Performance and Robustness of a Multistatic MIST Beamforming Algorithm for Breast Cancer Detection”, *Progress In Electromagnetics Research*, Vol. 105, pp. 403-424, 2010.
86. J. Capon, “High-Resolution Frequency-Wavenumber Spectrum Analysis”, *Proceedings of the IEEE*, Vol. 57, No. 8, pp. 1408-1418, 1969.
87. J. Li, P. Stoica and Z. Wang, “On Robust Capon Beamforming and Diagonal Loading”, *IEEE Transactions on Signal Processing*, Vol. 51, No. 7, pp. 1702-1715, 2003.
88. J. Li and P. Stoica, *Robust Adaptive Beamforming*. John Wiley & Sons. NJ, 2006.
89. R. G. Lorenz and S. P. Boyd, “Robust Minimum Variance Beamforming”, *Proceedings Asilomar Conference on Signals, Systems and Computers*, Vol. 2, pp. 1345-1352, 2003.
90. R. G. Lorenz and S. P. Boyd, “Robust Minimum Variance Beamforming”, *IEEE Transactions on Signal Processing*, Vol. 53, pp. 1684-1696, 2005.
91. B. Guo, Y. Wang, J. Li, P. Stoica and R. Wu, “Microwave Imaging Via Adaptive Beamforming Methods for Breast Cancer Detection”, *Progress In Electromagnetics Research Symposium*, Hangzhou, China, pp. 350-353, 2005.
92. B. Guo, Y. Wang and J. Li, “Microwave Imaging Via Adaptive Beamforming

- Methods for Breast Cancer Detection”, *Journal of Electromagnetic Waves and Applications*, Vol. 20, No. 1, pp. 53-63, 2006.
93. Y. Xie, B. Guo, L. Xu, J. Li and P. Stoica, “Multistatic Adaptive Microwave Imaging for Early Breast Cancer Detection”, *IEEE Transactions on Biomedical Engineering*, Vol. 53, No. 8, pp. 1647-1657, 2006.
94. Y. Xie, B. Guo, J. Li and P. Stoica, “Novel Multistatic Adaptive Microwave Imaging Methods for Early Breast Cancer Detection”, *EURASIP Journal on Applied Signal Processing*, Vol. 2006, pp. 1-13, 2006.
95. Y. Chen, E. Gunawan, K. S. Low, S. Wang, C. B. Soh and J. Lavanya, “Effect of Lesion Morphology on Microwave Signature in Ultra-Wideband Breast Imaging: A Preliminary Two-dimensional Investigation”, *Antennas and Propagation Society International Symposium, 2007 IEEE*, pp. 2168-2171, 2007.
96. Y. Huo, R. Bansal and Q. Zhu, “Breast Tumor Characterization Via Complex Natural Resonances”, *IEEE MTT-S International Microwave Symposium Digest*, Philadelphia, Vol. 1, pp. 387-390, 2003.
97. Y. Huo, R. Bansal and Q. Zhu, “Modeling of Noninvasive Microwave Characterization of Breast Tumors”, *IEEE Transactions on Biomedical Engineering*, Vol. 51, No. 7, pp. 387-390, 2004.
98. Y. Chen, E. Gunawan, K. S. Low, S. Wang, C. B. Soh and T. C. Putti, “Effect of Lesion Morphology on Microwave Signature in 2-D Ultra-Wideband Breast Imaging”, *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 8, pp. 2011-2021, 2008.
99. A. Mashal, B. Sitharaman, J. H. Booske and S. C. Hagness, “Dielectric Characterization of Carbon Nanotube Contrast Agents for Microwave Breast Cancer Detection”, *IEEE Antennas and Propagation Society International Symposium*, Charleston, South Carolina, USA, pp. 1-4, 2009.
100. A. Mashal, J. H. Booske and S. C. Hagness, “Toward Contrast-Enhanced Microwave-Induced Thermoacoustic Imaging of Breast Cancer: An Experimental Study of the Effects of Microbubbles on Simple Thermoacoustic Targets”, *Physics in Medicine and Biology*, Vol. 54, pp. 641-650, 2009.
101. K. Muinonen, “Introducing the Gaussian Shape Hypothesis for Asteroids and Comets”, *Astronomy and Astrophysics*, Vol. 332, pp. 1087-1098, 1998.
102. K. Muinonen, “Chapter 11: Light Scattering by Stochastically Shaped

- Particles”, in *Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications*, Editor(s): M. I. Mishchenko, J. W. Hovenier and L. D. Travis. 2000, Academic Press.
103. K. Fukunaga, *Introduction to Statistical Pattern Recognition*. Second Edition. M. Kaufmann. San Francisco, California, USA, 1990.
 104. K. Muinonen. *Gaussian Random Sphere Program G-sphere*. 2002. Last accessed: 23/05/2008; Available from: www.astro.helsinki.fi/psr/.
 105. E. C. Fear, J. Sill and M. A. Stuchly, “Experimental Feasibility Study of Confocal Microwave Imaging for Breast Tumor Detection”, *IEEE Transactions on Microwave Theory and Techniques*, Vol. 51, No. 3, pp. 887-892, 2003.
 106. E. C. Fear and M. Okoniewski, “Confocal Microwave Imaging for Breast Tumor Detection: Application to a Hemispherical Breast Model”, *Microwave Symposium Digest, 2002 IEEE MTT-S International, 2002*, Seattle, WA, USA, Vol. 3, pp. 1759-1762, 2002.
 107. I. J. Craddock, M. Klemm, J. Leendertz, A. W. Preece and R. Benjamin, “Development and Application of a UWB Radar System for Breast Imaging”, *2008 Loughborough Antennas & Propagation Conference*, 2008.
 108. E. C. Fear, A. Low, J. Sill and M. A. Stuchly, “Microwave System for Breast Tumor Detection: Experimental Concept Evaluation”, *Antennas and Propagation Society International Symposium, 2002. IEEE*, Vol. 1, pp. 819-822, 2002.
 109. M. Klemm, I. J. Craddock, J. A. Leendertz, A. Preece and R. Benjamin, “Radar-Based Breast Cancer Detection Using a Hemispherical Antenna Array – Experimental Results”, *IEEE Transactions on Antennas and Propagation*, Vol. 57, No. 6, pp. 1692-1704, 2009.
 110. H. Wold, “Estimation of Principal Components and Related Models by Iterative Least Squares”, in *Multivariate Analysis*, Editor(s): K. R. Krishnaiah. 1996, Academic Press: New York. pp. 391-420.
 111. J. Shlens. *A Tutorial on Principal Component Analysis*. Last accessed: 05/07/2010; Available from: http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf.
 112. D. J. Bartholomew, F. Steele, I. Moustak and J. I. Galbraith, *The Analysis and Interpretation of Multivariate Data for Social Scientists*. Texts in Statistical

- Science. Chapman & Hall/CRC. USA, 2002.
113. A. Hyvärinen and E. Oja, “Independent Component Analysis: Algorithms and Applications”, *Neural Networks*, Vol. 13, No. 4-5, pp. 411-430, 2000.
 114. H. Gävert, J. Hurri, J. Särelä and A. Hyvärinen. *The FastICA Package for MATLAB*. Last accessed: 20/07/2010; Available from: www.cis.hut.fi/projects/ica/fastica/.
 115. N. Saito and R. R. Coifman, “Local Discriminant Bases and Their Applications”, *Journal of Mathematical Imaging and Vision*, Vol. 5, pp. 337-358, 1995.
 116. Y. Mallet, D. Coomans, J. Kautsky and O. De Vel, “Classification Using Adaptive Wavelets for Feature Extraction”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 10, pp. 1058-1066, 1997.
 117. S. G. Mallat, “A Theory for Multiresolution Signal Decomposition: The Wavelet Representation”, *IEEE Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674-693, 1989.
 118. C. Valens. *A Really Friendly Guide to Wavelets*. 1999-2004. Last accessed: 01/07/2010; Available from: <http://pagesperso-orange.fr/polyvalens/clemens/wavelets/wavelets.html#section1>.
 119. PyWavelets. *Coiflets 5 Wavelet (Coif5) Properties, Filters and Functions - Wavelet Properties Browser*. Last accessed: 30/06/2010; Available from: <http://wavelets.pybytes.com/wavelet/coif5>.
 120. C. Hsu, C. Chang and C. Lin. *A Practical Guide to Support Vector Classification*. Last accessed: 03/04/2010; Available from: www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.
 121. W. J. Krzanowski, *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press. New York, 1988.
 122. B. S. Everitt and G. Dunn, *Applied Multivariate Data Analysis*. Second Edition. Arnold Publishers. New York, 2001.
 123. G. A. F. Seber, *Multivariate Observations*. John Wiley & Sons, Inc. Hoboken, New Jersey, 1984.
 124. T. Raykov and G. A. Marcoulides, *An Introduction to Applied Multivariate Analysis*. Routledge Taylor & Francis Group. New York, 2008.
 125. J. F. Hair, W. C. Black, B. J. Babin, R. E. Anderson and R. L. Tatham, *Multivariate Data Analysis*. Sixth Edition. Prentice Hall. Upper Saddle River,

- New Jersey, USA, 2006.
126. B. E. Boser, I. M. Guyon and V. N. Vapnik, "A Training Algorithm for Optimal Margin Classifiers", *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, PA, USA, pp. 144-152, 1992.
 127. C. Cortes and V. Vapnik, "Support-Vector Networks", *Machine Learning*, Vol. 20, No. 3, pp. 273-297, 1995.
 128. A. Ng. *Support Vector Machines (Part V of CS229 Machine Learning Course Materials)*. Last accessed: 13/05/2010; Available from:
<http://www.stanford.edu/class/cs229/notes/cs229-notes3.pdf>.
 129. K. P. Bennett and C. Campbell, "Support Vector Machines: Hype or Hallelujah?", *ACM SIGKDD Explorations Newsletter*, Vol. 2, No. 2, pp. 1-13, 2000.
 130. C. Campbell. *Introduction to Support Vector Machines*. 2008. Last accessed: 05/02/2008; Available from:
http://videlectures.net/epsrws08_campbell_isvm/.

Appendix

Publications

The publications resulting from the work presented in this thesis are listed here. The URLs to the papers are included at the end of each reference.

A.1. Published Journal Papers

- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Comparison of Planar and Circular Antenna Configurations For Breast Cancer Detection Using Microwave Imaging", *Progress In Electromagnetics Research*, Vol. 99, pp. 1-20, 2009 [49]. (<http://www.jpier.org/PIER/pier99/01.09100204.pdf>)
- **R. C. Conceição**, M. O'Halloran, E. Jones and M. Glavin, "Investigation of Classifiers for Early-Stage Breast Cancer Based on Radar Target Signatures", *Progress In Electromagnetics Research*, Vol. 105, pp. 295-311, 2010 [50]. (<http://www.jpier.org/PIER/pier105/17.10051904.pdf>)
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Support Vector Machines for the Classification of Early-Stage Breast Cancer Based on Radar Target Signatures", *Progress In Electromagnetics Research B*, Vol. 23, pp. 311-327, 2010 [51]. (<http://www.jpier.org/PIERB/pierb23/18.10062407.pdf>)
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Evaluation of Features and Classifiers for Classification of Early-Stage Breast Cancer", *Journal of Electromagnetic Waves and Applications*, Vol. 25, pp. 1-14, 2011 [52]. (<http://www.ingentaconnect.com/content/vsp/jew/2011/00000025/00000001/art00001>)

- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Effects of Dielectric Heterogeneity in the Performance of Breast Tumour Classifiers", *Progress In Electromagnetics Research M*, Vol. 17, pp. 73-86, 2011 [53]. (<http://www.jpier.org/PIERM/pierm17/06.10122402.pdf>)
- M. O'Halloran, **R. C. Conceição**, D. Byrne, M. Glavin and E. Jones, "FDTD Modeling of the Breast: A Review", *Progress In Electromagnetics Research B*, Vol. 18, pp. 1-24, 2009 [54] (contribution in Section 4). (<http://www.jpier.org/PIERB/pierb18/01.09080505.pdf>)
- B. McGinley, M. O'Halloran, **R. C. Conceição**, F. Morgan, M. Glavin and E. Jones, "Spiking Neural Networks for Breast Cancer Classification Using Radar Target Signatures", *Progress In Electromagnetics Research C*, Vol. 17, pp. 79-94, 2010 [55] (contribution in Sections 2 and 4.1). (<http://www.jpier.org/PIERC/pierc17/07.10100202.pdf>)
- M. O'Halloran, B. McGinley, **R. C. Conceição**, F. Morgan, M. Glavin and E. Jones, "Spiking Neural Networks for Breast Cancer Classification in a Dielectrically Heterogeneous Breast", *Progress In Electromagnetics Research*, Vol. 113, pp. 413-428, 2011 [56] (contribution in Sections 2 and 4.1.1). (<http://www.jpier.org/PIER/pier113/24.10122203.pdf>)

A.2. Published Conference Papers

- **R. C. Conceição**, D. Byrne, M. O'Halloran, M. Glavin and E. Jones, "Classification of Suspicious Regions within Ultrawideband Radar Images of the Breast", *16th IET Irish Signals and Systems Conference, ISSC 2008*, Galway, Ireland, Vol. 1, pp. 60-65, 2008 [57]. Also presented at Research Day for College of Engineering, NUIG in April 2008 (awarded second best poster). (<http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=IEECPS0020080CP539000060000001&idtype=cvips&gifs=yes&ref=no>)
- **R. C. Conceição**, M. O'Halloran, M. Glavin and E. Jones, "Antenna

configurations for Ultra Wide Band Radar Detection of Breast Cancer”, in *Proceedings of the SPIE BIOS West*, San José, CA, USA, Vol. 7169, No. 9, pp. [71691M, 12], 2009 [58]. Also presented at Research Day for College of Engineering, NUIG in April 2010 (awarded best presentation with Economical Potential). (http://spie.org/x648.html?product_id=808253)

- **R. C. Conceição**, M. O'Halloran, D. Byrne, E. Jones and M. Glavin, “Tumor Classification Using Radar Target Signatures”, *Progress In Electromagnetics Research Symposium*, Cambridge, MA, USA, pp. 346-349, 2010 [59]. (<http://www.piers.org/piersproceedings/download.php?file=cGllcnMyMDEwQ2FtYnJpZGdlfDFQN18wMzQ2LnBkZnwwOTEyMTkwODA5MjU=>)