

Ultrasonic Estimation of Mean Scatterer Spacing and Effective Scatterer Diameter for Breast Lesion Classification

A thesis submitted to the Department of Electrical and Electronic Engineering
of
Bangladesh University of Engineering and Technology
in partial fulfillment of the requirement for the degree of
MASTER OF SCIENCE IN ELECTRICAL AND ELECTRONIC ENGINEERING

by
Navid Ibtehaj Nizam
Student ID: 1015062229 P



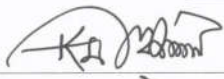



DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
BANGLADESH UNIVERSITY OF ENGINEERING AND TECHNOLOGY

April 2019

Approval Certificate

The thesis titled “**Ultrasonic Estimation of Mean Scatterer Spacing and Effective Scatterer Diameter for Breast Lesion Classification**”, submitted by Navid Ibtehaj Nizam, Student No.: 1015062229 P, Session: October, 2015, has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Master of Science in Electrical and Electronic Engineering on April 30, 2019.

Board of Examiners

1. 
 (Dr. Md. Kamrul Hasan)
 Professor
 Department of Electrical and Electronic Engineering
 Bangladesh University of Engineering and Technology
 Dhaka - 1205, Bangladesh
Chairman
 (Supervisor)
2. 
 (Dr. Md. Shafiqul Islam)
 Professor and Head
 Department of Electrical and Electronic Engineering
 Bangladesh University of Engineering and Technology
 Dhaka - 1205, Bangladesh
Member
 (Ex-Officio)
3. 
 (Dr. S. M. Mahbubur Rahman)
 Professor
 Department of Electrical and Electronic Engineering
 Bangladesh University of Engineering and Technology
 Dhaka - 1205, Bangladesh
Member
4. 
 (Dr. A. B. M. Aowlad Hossain)
 Professor
 Department of Electronics and Communication Engineering
 Khulna University of Engineering & Technology
 Khulna - 9203, Bangladesh
Member
 (External)

Candidate's Declaration

It is hereby declared that this thesis or any part of it has not been submitted elsewhere for the award of any degree or diploma.

Signature of the candidate

David

Navid Ibtehaj Nizam

(Student No.: 1015062229 P)

Dedication

To my mother and the memory of my late father.

Acknowledgements

First and foremost, I am grateful to the Almighty for giving me the opportunity and strength to carry out this thesis work.

I would like to extend my heart-felt gratitude to my supervisor, Dr. Md. Kamrul Hasan for giving me the opportunity to work on this topic. I am grateful for his unconditional support, constant guidance and supervision throughout this research work. Working under his supervision is a reminder of the discipline, patience, and perseverance required to carry out high-level research.

I am also grateful to the Head of the Department of Electrical and Electronic Engineering, BUET, for the research lab facilities. I would also like to thank the present and past members of the DSP research lab, without whose help this thesis may not have been possible, especially, Dr. Sharmin Rowshan Ara, Mr. Md. Rezaul Karim Tulib, Mr. Md. Hadiur Rahman Khan and Mr. Md. Shifat-E-Rabbi. Moreover, I would like to thank my colleagues at the Department of Biomedical Engineering, BUET, who were a constant source of support and inspiration and my friends, respected seniors, and beloved juniors at BUET, who are all shining examples to follow.

Furthermore, I would like to thank my relatives and in-laws, who are my pillars of strength and guidance. And I thank my parents for encouraging me all the time-my mother who always believes in me and my father, who was proud of even my tiniest achievements and whom, I wish, more than anything, was alive to see this day. Last but not the least, I would like to thank my loving wife, Nazia Afroz Choudhury, for all her sacrifice, help, and companionship.

This work has been supported by Higher Education Quality Enhancement Project, University Grants Commission (CPSF#96/BUET/Win-2/ST(EEE)/2017), Bangladesh. The *in vivo* breast data were acquired at BUET Medical Center by Dr. Farzana Alam, Assitant Professor, Department of Radiology and Imaging, Bangabandhu Sheikh Mujib Medical University, Dhaka-1000, Bangladesh.

Contents

Candidate's Declaration	i
Dedication	iii
Acknowledgements	iv
Abstract	xi
1 Introduction	1
1.1 Motivation of the Thesis	1
1.2 MSS and ESD Estimation Techniques and Ultrasonic Tissue Characterization: Literature Review	3
1.3 Objectives of the Thesis	5
1.4 Organization of the Thesis	6
2 Mean Scatterer Spacing Estimation	7
2.1 The Signal Model	7
2.2 Patient Data	8
2.3 FIELD II Simulator Setup	9
2.4 The MSS Estimation Technique	11
2.4.1 Preprocessing	11
2.4.2 EEMD Domain AR Spectrum Estimation	13
2.4.3 Order Selection Criterion for the EEMD Domain AR Model	14
2.4.4 MSS Estimation from the EEMD Domain AR Spectrum	16
2.5 Results on Simulation Phantom Data	17
2.6 Results on <i>in vivo</i> Breast Tissues	22
3 Effective Scatterer Diameter Estimation	24
3.1 Patient Data	24
3.2 TMP Data	24

3.3	The ESD Estimation Technique	25
3.3.1	Preprocessing	25
3.3.2	Proposed ESD Estimation Technique	26
3.4	Experimental Phantom Results	32
3.5	Results on <i>in vivo</i> Breast Tissues	35
4	Breast Lesion Classification	39
4.1	Classification Dataset	39
4.2	Micro- and Macro-Parameter Based Classification of Breast Lesions . .	40
4.3	Classification Techniques	40
4.4	Classification Results	43
5	Conclusion, Limitation and Future Scope	46
5.1	Conclusion	46
5.2	Limitation and Future Scope	47
	List of Publications	49

List of Figures

2.1	(a) A $40 \times 10 \times 6$ mm ³ phantom generated using the Field II software and (b) a B-mode image of the same phantom.	10
2.2	A block diagram illustrating our proposed algorithm.	12
2.3	Plot of MAPE against AR model order in EEMD domain for the deconvolved synthetic backscattered RF data generated using FIELD II. . . .	15
2.4	Comparison between different methods in estimating MSS values for simulated backscattered RF data. The methods are (a) conventional AR spectrum, (b) conventional AR cepstrum, (c) SSA, (d) SAC, (e) Modified SAC, (f) GS, and (g) proposed method.	18
2.5	Variation of mean SD of the estimated MSS for different methods as the simulated MSS values are gradually increased for SD of regular scatterer spacing of (a) 0 % and (b) 10%.	19
2.6	Effect of changing (a) the relative reflectivity of the regular and diffuse scatterers, (b) the consistency of spacing of the regular scatterers and (c) the phantom size for a simulated MSS value of 1 mm.	21
3.1	A block diagram illustrating our proposed algorithm. NNARLF refers to nearest neighborhood average regression line fitting	26
3.2	An illustration of the exponentially weighted neighborhood.	31
3.3	Scatter plot showing the variation of ESD across different windows on moving laterally from the edge of the ROI in the TMP background to the edge of the ROI within the inclusion.	34
3.4	Bar plot showing the impact on the percentage error and SD of the ESD estimates from the experimental TMPs on removing the different steps from our proposed algorithm.	35

3.5	Histology images of (a) fibroadenoma, (b) malignant and (c) inflammatory breast tissue. Scale bar: 100 μm in (a) and 200 μm in (b) and (c).	37
3.6	ESD map for four ROIs for a representative fibroadenoma tissue. A deeper shade of blue indicates higher ESD values. The values on the axes are arbitrary.	38

List of Tables

2.1	Simulation Parameters	11
2.2	Mean SD and MAPE obtained for different methods for MSS estimation from simulated backscattered RF data	17
2.3	Estimated values of MSS with SD (in bracket) using different methods for different tissue types	23
3.1	Description of Experimental TMPs	25
3.2	Estimated ESD values (in μm) with SD (in bracket) from experimental TMPs	33
3.3	Estimated values of ESD with SD (in bracket) using different methods for different tissue types	36
3.4	Estimated values of ESD with SD (in bracket) for normal tissues with different neighborhood sizes.	37
4.1	List of bi-modal QUS macro-paramters	41
4.2	Breast lesion classification results obtained for the different micro- and macro-parameters	44

Glossary

CAD	Computer-aided diagnosis
EEMD	Ensemble empirical mode decomposition
ESD	Effective scatterer diameter
MAPE	Mean absolute percentage error
MASD	Mean absolute squared deviation
MSS	Mean scatterer spacing
NN	Nearest neighbors
PSF	Point spread function
QUS	Quantitative ultrasound
RF	Radio-frequency
ROI	Region of interest
SD	Standard deviation
TMP	Tissue mimicking phantom
UB	Ultrasound B-mode
UE	Ultrasound elastography

Abstract

Quantitative ultrasound (QUS) based parameters, estimated from backscattered radio-frequency (RF) data, allows the parametrization of tissue micro-structures. These parameters, which include mean scatterer spacing (MSS) and effective scatterer diameter (ESD), often reveal more information about the interrogated tissue than conventional B-mode imaging, as well as being less subjective to operator settings and interpreter variability than conventional ultrasound. MSS and ESD are important QUS micro-parameters for detecting pathological changes in breast tissue. In this thesis, two novel techniques are proposed for estimation of MSS and ESD from breast tissues. Both these techniques rely on the separation of the coherent and diffuse component of backscattered data using ensemble empirical mode decomposition (EEMD) of the data into their intrinsic mode functions (IMFs). An automatic IMF selection scheme is employed, which utilizes a non-parametric Kolmogorov-Smirnov (K-S) test to automatically select the IMFs responsible for coherent scattering in case of MSS estimation, and diffuse scattering in case of ESD estimation. Before EEMD can be performed, filtering and deconvolution of the backscattered data is carried out to reduce the impacts of diffraction and the system point spread function (PSF). The MSS is estimated from the spacing between the peaks of the power spectrum estimated from the coherent component of RF data. The power spectrum is estimated using an autoregressive (AR) model, whose order is chosen by minimization of a novel mean absolute percentage error (MAPE) criterion. The ESD is estimated from the diffuse component of RF data utilizing a theoretical tissue scattering model in the frequency domain. MSS estimation is carried out on simulation RF data generated by FIELD II and *in vivo* breast tissues while ESD estimation is carried out on experimental tissue-mimicking phantoms (TMPs) and *in vivo* tissues. The average MSS for normal tissues, inflammatory tissues, fibroadenoma, and malignant tissues are found to be 0.689 (± 0.032) mm, 0.729 (± 0.040) mm, 0.750 (± 0.035) mm, and 0.793 (± 0.040) mm, respectively. The corresponding average ESD values are 75.12 (± 4.01) μm , 75.72 (± 4.09) μm , 98.71

(± 9.55) μm , and 123.05 (± 8.85) μm , respectively. The estimated average MSS and ESD values correspond well to those previously reported in literature. When MSS and ESD are combined with 27 previously reported QUS bi-modal (ultrasound B-mode (UB) and ultrasound elastography (UE)) macro-parameters, to form a unique hybrid micro-macro feature set, consisting of 29 parameters, for binary (benign-malignant) classification of breast lesions, we obtain sensitivity, specificity, and accuracy, values as high as 98.21%, 98.06%, and 98.11%, respectively, using machine learning algorithms. This highlights the potential of this hybrid feature set as a computer-aided diagnosis (CAD) tool for breast lesion classification.

Chapter 1

Introduction

In this Chapter, we discuss the motivation behind the development of novel techniques for mean scatterer spacing (MSS) and effective scatterer diameter (ESD) estimation, and developing a hybrid feature set based breast lesion classification system. A detailed review of the relevant techniques for MSS and ESD estimation and QUS based breast lesion classification results reported in the literature is also presented. Next, the primary objectives of this thesis are highlighted. Finally, the organization of the thesis is described.

1.1 Motivation of the Thesis

Breast cancer is one of the leading causes of death in women all over the world [1]. The only way to improve the chances of survival is through early detection [2]. Breast cancer diagnosis, prognosis, and treatment require identification of breast lesion pathologic characteristics. For a long time, only biopsy, an invasive method, had been used for such characterization [3]. Recently, noninvasive quantitative ultrasound (QUS) methods have become increasingly popular for the characterization of breast lesions [4], [5]. QUS has been employed for the parametrization of tissue micro-structures with a view to developing a diagnostic modality which is less subjective to operator settings and interpreter variability compared to conventional ultrasound imaging [6]. The tissue characterization methods based on QUS parameters mainly rely on the concept that

disease processes alter the physical properties of tissues and hence, their acoustic scattering properties [7]. QUS parameters are being used in computer-aided diagnosis (CAD) tools for classifying between benign and malignant lesions [8]. MSS and ESD are important QUS based micro-parameters for breast lesion classification [5]. Although, quite a few techniques have been developed for MSS and ESD estimation [4], [9], [10], very few algorithms have been developed for MSS and ESD estimation of breast tissues, due to the complexity of the breast structure [11]. Ultrasound scattering from the female breast has two principal components. One component, which is highly coherent, is due to scatterers that are spaced at relatively regular intervals and the other one is due to random or diffuse scatterers [4], [12]. Micro-parameters like MSS are estimated from the coherent component from backscattered data while others, such as ESD, are estimated from the diffuse component [4], [13]. Hence, successful separation of these two components is essential in any algorithm for micro-parameter estimation [14], [15]. Although some techniques have been developed to separate the diffuse component from the coherent component of backscattered data [13], [14], these techniques are highly sensitive to noise and remain untested on breast tissue. Therefore, in this work, we have attempted to develop an ensemble empirical mode decomposition (EEMD) based signal decomposition technique for MSS and ESD estimation of female breast tissues. The inherent advantages of EEMD in that it requires no pre-selection of basis function [16], is completely data-driven [16], and that it works equally well on both stationary and non-stationary signals [17], making it suitable for separating the coherent component and the diffuse component of backscattered data. Furthermore, to the best of our knowledge, no previous work has been reported on a broad dataset for classifying between benign and malignant breast lesions using ultrasonic micro-parameters like MSS and ESD, either individually or combined. Additionally, another class of QUS parameters, usually extracted from the ultrasound B-mode (UB) images and ultrasound elastography (UE), referred to as macro-parameters have also been successfully employed for breast lesion classification [18]. However, no previously reported technique has attempted to combine the micro- and macro-parameters for ultrasonic tissue characterization. Thus, we have looked to combine the 2 micro-parameters, MSS and ESD,

and 27 previously reported ultrasonic bi-modal (UB and UE) macro-parameters [18], to form a hybrid feature set for binary (benign-malignant) classification of breast lesions. The motivation behind combining these two unique feature sets is that the micro-parameters encompass the information relating to tissue histopathology while the macro-parameters capture the information relating to radiological imaging and hence, combining these features to train a machine learning network should allow for better classification of breast lesions.

1.2 MSS and ESD Estimation Techniques and Ultrasonic Tissue Characterization: Literature Review

Fellingham first proposed the use of MSS in 1979 [19] for the classification of tissues, and later, Fellingham and Sommer [20] used this parameter to differentiate between normal and cirrhotic liver. They used a periodogram approach to compute the fast Fourier transform (FFT) which, depending on the window selected, can lead to a degradation of the resolution. To overcome this problem, an AR model-based spectral estimation scheme is often used [21]. But this method does not take into consideration the effect of the transducer response embedded in the RF data [4]. Hence, Kuc *et al.* [22] suggested the use of the magnitude of the cepstrum instead of the AR spectrum to estimate the MSS. Pereira *et al.* [14] used singular spectrum analysis (SSA) to estimate MSS values for experimental phantom and bovine liver which uses some heuristic criteria for the proper selection of eigenvectors. Georgiou *et al.* [11] used a wavelet-based decomposition method to decompose a backscattered RF signal into its coherent and diffuse components and used this to extract features from breast RF data for classification. Rosado-Mendez *et al.* [23] used an empirical method to estimate the MSS from tissue mimicking phantoms and cervix tissues. The method, however, is highly dependent on the size of the resolution cell. Although the literature is laden with studies on estimating the MSS of tissue mimicking phantoms and liver tissues, work on estimating the MSS of breast tissues is very limited. The estimation of MSS from breast tissues is challenging because of the irregularity of the breast structure

compared to the liver [24]. Donohue *et al.* [25] introduced the generalized spectrum (GS) to analyze the scatterer spacing of breast tissues and was later successful in classifying breast tissues by extracting parameters from the collapsed average (CA) of the GS using a linear discriminant analysis (LDA) [26]. Rubert *et al.* [27] expanded the use of the GS for estimating the MSS of liver tissues using a multi-taper spectral estimation method. Bige *et al.* [4] used an AR cepstrum method to estimate the MSS of breast tissues. However, this method is not immune to interference from diffuse scatterers since it makes no attempt to recover the fundamental harmonic from the backscattered signal or improve its periodicity. Varghese *et al.* [15] suggested the use of a spectral autocorrelation (SAC) function to estimate the MSS of liver tissues which was later modified by Taddayon *et al.* [5] for identification and classification of pathological changes in breast tissues. In computing the SAC, the position of the diffuse and the coherent scatterers are assumed to be uncorrelated while the diffuse scatterer process is assumed to be wide-sense stationary over the region of interest. However, any significant deviation from these assumptions may lead to incorrect MSS estimates for *in vivo* data.

The use of ESD, as a micro-parameter for characterizing pathological tissues, has also been reported extensively in the literature. These include the ESD of ocular lesions [28], liver [29], renal tissues [30], glomerular tissues [31], kidney [32], prostate [33], human aortae [34], and the uveal melanomas [35]. In addition, ESD has been successfully employed to distinguish between benign and malignant lesions in the eyes [36] and in the lymph nodes [37]. Ultrasonic characterization of human breast tissues have been reported in [5] in [38]. However, in the method discussed in [38], no specific micro-parameters have been used while [5] used micro-parameters for tumor grading rather than breast lesion classification. As stated previously, no previous work has been reported on a large dataset, for classifying between benign and malignant breast lesions using ultrasonic micro-parameters like MSS and ESD. In [39], it has been reported that, for the estimation of ESD, the coherent component behaves as an interference and hence, needs to be suppressed. The generalized spectrum and Rayleigh envelop statistics [40] and Hanning tapers [41] have been used to separate the diffuse echoes

from the backscattered data but these algorithms remain untested on human tissue. Furthermore, the existing techniques use large 2-D spatial signal blocks to generate a stable block power spectra [10], [42] which is often an unrealistic approach because the tissue pathology inside a large spatial region must be considered uniform, a requirement that is too idealistic for a heterogeneous tissue medium. Additionally, most ESD estimation techniques use a form factor to model the backscattering from tissue structures [10]. A mean average squared deviation (MASD) minimization technique has been proposed in [10] while a frequency domain technique has been proposed in [9], which employ a Gaussian form factor to model tissue scattering. It has been previously established that the Gaussian form factor most accurately models ultrasonic scattering from human tissues [10], [43]. What is more, as discussed before, macro-parameters derived from the UB image and UE have been successfully employed for classification of breast lesions [18], [44]. However, no previous method exists that has tried to combine the micro- and macro-parameters for ultrasonic tissue characterization.

1.3 Objectives of the Thesis

The objectives of this work are:

1. To propose new techniques for MSS and ESD estimation of breast tissues based on the decomposition of ultrasound RF echo signals.
2. To combine MSS and ESD with other QUS based macro-parameters to develop a robust computer-aided diagnosis (CAD) tool for breast lesion classification using machine learning.
3. To design an ultrasound RF simulator, using FIELD II, that can generate RF data corresponding to different scattering conditions.
4. To make a comparative performance study of the proposed method with other reported MSS and ESD estimation techniques using known ESD experimental phantom data, known MSS simulation RF data, and in vivo breast data.

1.4 Organization of the Thesis

This thesis consists of five chapters. Chapter 1 is composed of the motivation and objectives of the thesis and the limitations of the existing techniques. In Chapter 2, the techniques for MSS estimation, the FIELD II simulator setup for MSS estimation from simulation data, the results of MSS estimation from *in vivo* breast tissues, and comparisons with some existing techniques are discussed under relevant sections. Chapter 3 has the detailed description of the theoretical aspects of the ESD estimation technique, presentation of the results obtained from experimental tissue-mimicking phantoms (TMPs) and *in vivo* breast tissues with relevant discussion, and comparisons with some existing techniques. In Chapter 4, we elaborate on the techniques employed for binary (benign-malignant) classification of breast lesions, and also present and discuss the obtained classification results. Finally, in Chapter 5, concluding remarks and suggestions for future research are provided based on the outcomes and limitations of this thesis work.

Chapter 2

Mean Scatterer Spacing Estimation

In this Chapter, we discuss the basic signal model and the rationale behind a signal decomposition based technique. We give the details of the patient data and the simulation data, with known MSS values, generated by the FIELD II software. We also discuss how the FIELD II software is used to simulate different scattering conditions. Next, we focus on the theoretical aspects of the MSS estimation technique and show that the simulation data can be used to verify the reliability of the technique under different scattering conditions and compare with some existing techniques, namely the conventional AR spectrum [45], the conventional AR cepstrum [4], singular spectrum analysis (SSA) [14], spectral autocorrelation (SAC) [15], modified SAC [5], and the GS [25]. In addition, we report the estimated average MSS values of different types of female breast tissues.

2.1 The Signal Model

The backscattered RF data are known to consist of two principal components. The component that is important for MSS estimation is the one due to the presence of regular or periodic scatterers along the ultrasound scan line. An interfering component is present due to the diffuse or random scatterers. The coherent scatterers of the breast can be considered as ducts or lobules, and diffuse scatterers as breast tissue cells [46]. Moreover, the echo signal produced by the system scatterers is convolved

with the system PSF to produce the measured RF signal. If $x_i(n)$ is taken to be the backscattered RF signal of the i -th scan line, it can be modeled as

$$x_i(n) = \left\{ \sum_{k=1}^{N_s} a_{ki}(n - \tau_{ki}) + \sum_{k=1}^{N_D} v_{ki}(n - \theta_{ki}) \right\} * h(n) \quad (2.1)$$

where n , N_D , v_{ki} , θ_{ki} , N_S , a_{ki} , and τ_{ki} represent the time index, the total number of diffuse scatterers, reflectivity of the k -th diffuse scatterer, the delay associated with the k -th diffuse scatterer, the total number of regular scatterers, the reflectivity of the k -th regular scatterer, and the delay associated with the k -th regular scatterer, respectively [15]. Here, $h(n)$ represents the system PSF. The scan lines of each 2D frame of the backscattered RF data acquired by the ultrasound imaging scanner (transducer) can be modeled, using Born approximation [47], [48], as outputs from a single-input multiple-output (SIMO) system with the ultrasound pulse transmitted by the transducer as the single input [49]. Hence, an 1D PSF, $h(n)$, as shown in (2.1), is usually used to model the system effects along each scan line [47], [48]. The resulting 1D PSF can be assumed to be proportional to a 1D Gaussian given by

$$h(z, n) \propto \frac{1}{w_z(2\pi)^{\frac{3}{2}}} \exp \left[-\frac{(z - \frac{wn}{2k})^2}{w_z^2} \right] \quad (2.2)$$

where w_z and k represent the standard deviation (SD) width in the z (axial) direction of the resolution cell and the wave number, respectively [50].

The signal model provides the motivation for our proposed method. If the system PSF is removed by deconvolution, then the regular component and the diffuse component become additive. This allows better separation of the regular component and the diffuse component by ensemble empirical mode decomposition (EEMD).

2.2 Patient Data

The *in vivo* data used in this thesis have been obtained at the Bangladesh University of Engineering and Technology (BUET) Medical Center with the help of a SonixTOUCH Research (Ultrasonix Medical Corporation, Richmond BC, Canada) scanner integrated

with a L14-5/38 linear probe. The probe was operating at 10 MHz with 65% bandwidth (-6-dB bandwidth) at a sampling rate of 40 MHz. The pulse length was approximately 0.4 mm and the beam width was approximately 2.2 mm. The pulse length has been estimated from the emitted pulse using the multiple input-output inverse theorem [51]. The study has been conducted on 179 patients, with their prior written consent, and approved by the institutional review board (IRB) of BUET. From the 179 patients, 245 RF data files have been recorded. Out of the 245 data records, 56 are malignant lesions, 79 are fibroadenomas, 24 are inflammatory growths, 42 are cystic lesions and 44 data records do not show any growths. The age range of all patients was 13-75 years (mean: 35.27 years). The patients having masses underwent fine-needle aspiration cytology (FNAC) and/or excision biopsy according to the suggestion of their physicians. All patients having FNAC diagnosis positive for malignancy underwent surgery. Therefore, diagnoses of malignant and some benign lesions were confirmed by histopathology, and diagnoses of the remaining lesions, by cytopathology.

2.3 FIELD II Simulator Setup

Numerical phantoms are generated using the FIELD II software [52] where the MSS is varied from 0.6 mm to 1.4 mm (with increment in steps of 0.1 mm). In the simulation, we have used a transducer center frequency of 10 MHz and a bandwidth of 7 MHz. The transmit focus is placed 60 mm from the transducer face. The pulse length and the beam width are arbitrarily chosen to be approximately 0.50 mm and 2.5 mm, respectively. The size of the ROI is chosen such that it is approximately 20 pulse widths axially and 4 beam widths laterally. The phantom consisted of both randomly generated scatterers (i.e., diffuse scatterers) and regular scatterers. The spacing between the regular scatterers are obtained from a gamma distribution with a standard deviation (SD) of 3% of mean spacing. The diffuse scatterers are uniformly distributed throughout the beam field with the number of scatterers generated using the Poisson distribution so that each resolution cell contains 15 diffuse scatterers. The FIELD II software performed the task of generating the backscattered RF data from the simula-

tion phantom. The RF data is generated by the convolution of the echo signal produced by the system scatterers with the system PSF. Table 2.1 provides a brief summary of the simulation parameters. A typical phantom generated using the Field II software is shown in Fig. 2.1(a). The scatterers are marked with black dots in the figure. A B-mode image of a plane of the same phantom generated using Field II is shown in Fig. 2.1(b). The B-mode imaging plane is sagittal with the transducer face placed at the bottom surface of the phantom shown in Fig. 2.1(a). The distance between the transducer face and the surface of the numerical phantom is 30 mm.

Moreover, analyses are performed on simulation data to observe the effect of varying

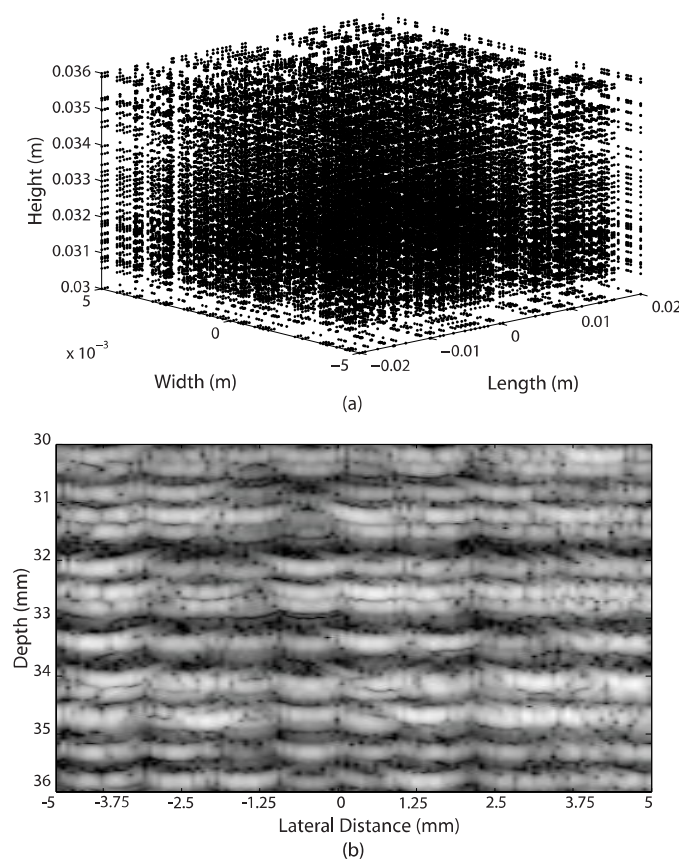


Figure 2.1: (a) A $40 \times 10 \times 6 \text{ mm}^3$ phantom generated using the Field II software and (b) a B-mode image of the same phantom.

the relative reflectivity of the regular and diffuse scatterers, varying the SD of the mean spacing of the regular scatterers and varying the axial size of the simulation phantom.

The relative reflectivity is varied by setting the scattering amplitude of the regular scatterers at 0.01, 0.1, 0.5, 1.0, 10, 50, and 100 times the scattering amplitude of the diffuse scatterers using Field II. The simulated MSS value is kept constant at 1 mm and the simulation parameters defined in Table 2.1 are used. Furthermore, the inconsistency of the spacing is varied by changing the SD of the mean of the gamma distributed spacing from 0% to 10% (in steps of 1%). The simulated MSS value is again set at 1 mm and the other simulation parameters are kept the same as those defined in Table 2.1. In order to investigate the effect of phantom size, and hence, attenuation on the accuracy of the different methods, the height of the simulation phantom is varied from 0.5 cm to 4 cm. A frequency dependent attenuation coefficient value of 0.5 dB/MHz/cm is set on Field II. This attenuation coefficient value is typical for soft tissues [53], [54]. As before, the simulated MSS value is set at 1 mm and the simulation parameters defined in Table 2.1 are used.

Table 2.1: Simulation Parameters

Number of diffuse scatterers per resolution cell	15 using a Poisson Distribution
Diffuse scatterer spatial distribution	Uniform
Regular scatterer spatial distribution	Gamma (SD of mean spacing of 3%)
Transducer Center Frequency	10 MHz
Bandwidth	7 MHz
Transmit Focus	60 mm from transducer face

2.4 The MSS Estimation Technique

2.4.1 Preprocessing

A block diagram of our proposed MSS estimation algorithm is shown in Fig. 2.2. The backscattered RF signal is stored in the form of a 3D video file, where the third

dimension is time. A 2D ROI is selected from a suitable frame of the acquired RF data. For selecting the ROI from a patient with a breast lesion, the B-mode image is observed, and a suitable 2D region within the border of the lesion is selected. Additionally, a second ROI is selected outside the lesion to compare the MSS values inside and outside the lesion. For a normal breast, a well-defined 2D region is selected, away from the edges of the imaging plane. The dimension of each ROI is taken to be approximately $(10 - 12) \times (6.25 - 9.40)$ mm². The beam width is approximately 2.2 mm. Therefore, each ROI contains approximately 100 resolution cells. The size of the ROI is approximately 25 pulse lengths axially and 4 beam widths laterally. Before

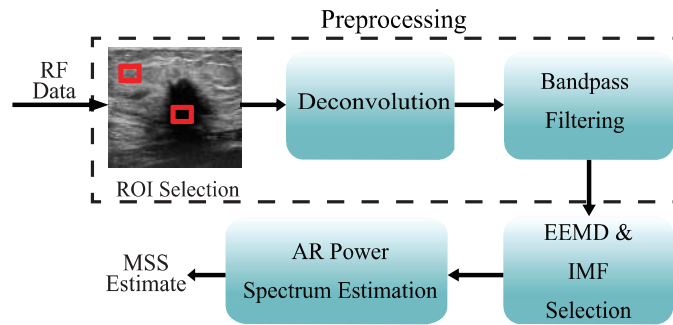


Figure 2.2: A block diagram illustrating our proposed algorithm.

EEMD domain analysis for MSS estimation, it is necessary that the effects of the system PSF are removed through a process called deconvolution and the impacts of diffraction are reduced. Applying a simple high-pass filter on the raw RF data to remove the low frequency system effects will not be effective since it will also significantly reduce the number of harmonics in the AR spectrum due to the regular scatterers in the low frequency region. In this work, a block-based damped variable step-size multichannel least mean square (bMCLMS) algorithm is used for blind deconvolution of the RF data [55]. It requires no prior knowledge of the PSF. Moreover, a modified block-based cross-correlation technique is used to overcome the non-stationarity of PSF and incomplete acquisition of the ultrasound data. Furthermore, an l_1 -norm based cost function is used along with a damped variable step-size. The damped variable step-size helps mitigate the effects of noise and results in a high convergence

speed. Finally, the PSF is estimated from lateral blocks of data using a regularized multiple-input/multiple-output algorithm, which works equally well for minimum and non minimum phase signals. This deconvolution method has been shown to outperform most conventional deconvolution techniques [55] and is, therefore, well suited for our work. After that, an ideal bandpass filter of frequency range 2 – 13 MHz is applied on the deconvolved data. The rationale behind choosing this frequency range is to reduce the impact of diffraction prominent below 2 MHz and eliminate other acquisition noises present at higher frequencies [56]. It is to be noted that deconvolution shifts the backscattered signal from being centered at the pulse frequency (10 MHz) to lower frequencies.

2.4.2 EEMD Domain AR Spectrum Estimation

The regular component of the backscattered signal is often masked by interference from the diffuse scatterers. Therefore, decomposition of the deconvolved data is required to separate the signal components due to regular scatterers and diffuse scatterers. Empirical mode decomposition (EMD), developed by Huang et al. [16] decomposes a signal into a sum of its IMFs. These IMFs are known to contain information about the different frequency components present in the signal. EMD is a plausible option for extracting the periodic component from the backscattered RF data as it does not require any pre-selection of the basis function, and rather, is a data-driven approach which is particularly suitable for non-stationary signals [16]. Using EMD, the deconvolved signal, $x_{id}(n)$, would be decomposed into the sum of its IMFs, $c_j(n)$, $j = 1, \dots, K$, where K is the number of IMFs and a residue $r(n)$ given by

$$x_{id}(n) = \sum_{j=1}^K c_j(n) + r(n). \quad (2.3)$$

However, in EMD, even a small perturbation can lead to a completely different set of IMFs, which may lead to a large variance in the MSS estimates [57]. In order to produce stable IMF estimates, EEMD is performed [17], [58], [59]. In EEMD [58], an ensemble of N_E random Gaussian noise, $w_{ip}(n)$ ($p = 1, \dots, N_E$), maintaining a signal-to-noise ratio (SNR) of 30 dB with the deconvolved signal, is added to the deconvolved

backscattered signal for the i -th scan line. That is, the ensemble is given by

$$x_{idp}(n) = x_{id}(n) + w_{ip}(n), p = 1, \dots, N_E. \quad (2.4)$$

The EMD algorithm is applied to each of the signals to extract the IMFs $c_{pj}(n)$ ($j = 1, \dots, K$) so that

$$x_{idp}(n) = \sum_{j=1}^K c_{pj}(n) + r_{pK}(n), p = 1, \dots, N_E \quad (2.5)$$

where $r_{pK}(n)$ is the residual function of $x_{idp}(n)$. Finally, the IMFs using EEMD are obtained from the ensemble average

$$\bar{c}_j(n) = \frac{1}{N_E} \sum_{p=1}^{N_E} c_{pj}(n), j = 1, \dots, K. \quad (2.6)$$

Note that, EEMD retains the effectiveness of EMD for the analysis of non-stationary signals [58]. As stated before, EEMD allows separation of the diffuse and the regular components from the deconvolved data and thereby, helps in reducing the effect of the diffuse scatterers on the regular scatterer components.

Selection of IMFs is a crucial issue in any algorithm involving EEMD [60], [61]. This is because, a few of the IMFs will contain information about the coherent scatterers while the others will contain information about the diffuse scatterers [60]. In order to identify the IMFs exhibiting coherent scattering, a non-parametric Kolomogorv-Smirnov (K-S) test is performed on the IMFs estimated using EEMD. The method used is described in [62]. The K-S classifier assumes that diffuse scatterers generate Gaussian statistics and any deviation is a result of coherent scattering. Those IMFs that do not show deviation from Gaussian statistics (at 5% significance level) are excluded and the remaining IMFs are summed up and used for further processing.

2.4.3 Order Selection Criterion for the EEMD Domain AR Model

The appropriate selection of the AR model order, p , is very important for accurate estimation of the MSS [63]. Bige *et al.* [4] suggested an AR model of order 90 that,

in their work, yielded the best results. However, using such a high AR model order is susceptible to erroneous results [63]. Too low an order does not show sufficient peaks in the spectrum. In order to choose the appropriate model order, a simulation was carried out using the FIELD II software [52] for synthetic values of MSS between 0.6 mm to 1.4 mm. The AR model order was varied from 30 to 100 and the AR spectrum and MSS were estimated using our proposed algorithm. The mean absolute percentage error (MAPE) was calculated for each model order. Fig. 2.3 shows a graph for MAPE against AR model order in the EEMD domain. We see from the figure that MAPE is lower around an AR model order of 60 ± 3 . Therefore, we chose an AR model order of 60 in our work. The AR model coefficients are estimated using the Burg's algorithm since it is a high resolution spectral estimation technique and results in a stable AR model [45]. It is also computationally efficient [45].

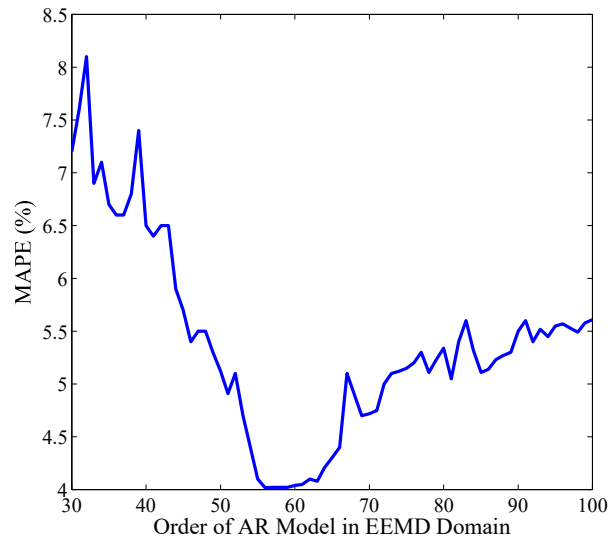


Figure 2.3: Plot of MAPE against AR model order in EEMD domain for the deconvolved synthetic backscattered RF data generated using FIELD II.

2.4.4 MSS Estimation from the EEMD Domain AR Spectrum

The MSS, d_i , for the i -th scan line, can be estimated from the separation between adjacent harmonics, Δf , present in the calculated EEMD domain AR spectrum for that scan line as

$$d_i = \frac{c}{2\Delta f} \quad (2.7)$$

where c is the speed of ultrasound waves in biological tissues [64]. If we estimate MSS from the separation between adjacent peaks, we obtain L_i MSS values d_{im} ($m = 1, 2, \dots, L_i$) where L_i is the number of peaks due to the regular scatterers in the i -th scan line and d_{im} are the L_i MSS values estimated from the separation between adjacent peaks. These L_i MSS values are averaged to obtain an estimate for the i -th scan line as

$$d_i = \frac{1}{L_i} \sum_{m=1}^{L_i} d_{im}. \quad (2.8)$$

After observing the EEMD domain spectra of several frames, it was found that peaks having amplitude less than 10% of the amplitude of the highest peak in the spectra can be treated as non-harmonic peaks. This procedure is repeated for each of the scan lines in the ROI and an estimate of the MSS, d , is taken to be the average of the MSS estimated from each of the scan lines given by

$$d = \frac{1}{N} \sum_{i=1}^N d_i \quad (2.9)$$

where N refers to the number of scan lines in the ROI. However, the MSS can be also estimated from the position of the fundamental harmonic obtained from the EEMD domain AR spectra of the i -th scan line, f_o^i , alone. In that case, the MSS for the i -th scan line is given by

$$d_i = \frac{c}{f_o^i}. \quad (2.10)$$

An estimate of the MSS, d , is taken to be the average of the MSS estimated from each of the scan lines given by (2.9). While estimating the MSS from a particular scan line, a check is performed to ensure that the L_i MSS values obtained from that scan line

have a standard deviation (SD) of less than 0.02 mm and there are at least five values of L_i , that is the AR spectrum of the frame must contain at least five harmonics due to the regular scatterers. If these criteria are not satisfied, that particular scan line is not included in the calculation of the overall MSS in (2.9). This helps to eliminate scan lines which may produce sporadic peaks in the AR spectrum.

2.5 Results on Simulation Phantom Data

For each MSS value, 10 data sets, each having a different spatial configuration, are generated. The MSS is estimated for all the 10 data sets using a particular method. The calculated MSS estimates are then averaged to obtain the estimate for a particular MSS value for a particular method. The MAPE and SD are also obtained for each of the simulated MSS values for a particular method. This procedure is repeated for all the simulated MSS values (from 0.6 mm to 1.4 mm) and the obtained SD values are averaged to obtain the mean SD. The MAPE is calculated in a similar manner. The results are summarized in Table 2.2.

Table 2.2: Mean SD and MAPE obtained for different methods for MSS estimation from simulated backscattered RF data

Method	MAPE (%)	Mean SD
AR spectrum [45]	27.76	± 0.18
AR cepstrum [4]	11.09	± 0.08
SSA [14]	6.03	± 0.07
SAC [15]	5.97	± 0.07
Modified SAC [5]	5.99	± 0.05
GS [25]	6.02	± 0.06
Proposed Method	5.78	± 0.06

Our proposed method results in a MAPE value of 5.78% which is significantly lower than the conventional AR spectrum and AR cepstrum methods and slightly lower than

SAC, modified SAC, GS, and SSA. The lowest mean SD of 0.05 is produced by the modified SAC method. However, it is to be noted that the MAPE of our proposed method is 0.21% lower than the modified SAC method. The estimated results are also illustrated graphically in Figs. 2.4(a)-(g). Each simulated MSS value (referred to as expected MSS in the horizontal axis) is generated for 10 different spatial configurations of coherent and diffuse scatterers. The corresponding values on the vertical axis refers to the MSS estimated by different methods for the simulated MSS values. It is clear from Figs. 2.4(a)-(g) that our proposed method has the lowest MAPE values overall for the entire MSS range of the simulation. The SAC, GS, SSA, and modified SAC

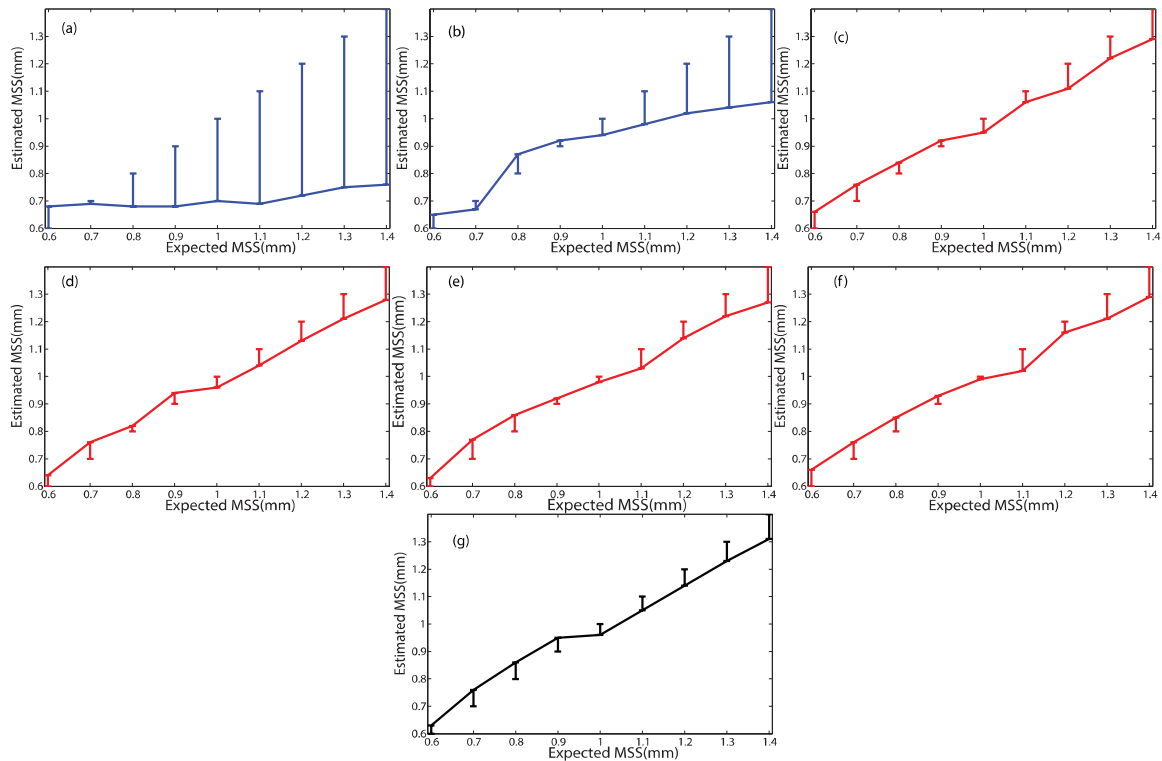


Figure 2.4: Comparison between different methods in estimating MSS values for simulated backscattered RF data. The methods are (a) conventional AR spectrum, (b) conventional AR cepstrum, (c) SSA, (d) SAC, (e) Modified SAC, (f) GS, and (g) proposed method.

have MAPE values close to our proposed method for the nearly deterministic scatterer

spacing. In order to better understand how the precision of these methods vary when the simulated MSS values are increased, plots have been shown in Figs. 2.5(a) and 2.5(b) to illustrate the change in SD as the expected MSS values are increased in simulation for a perfectly deterministic spacing (SD of 0%) and a highly irregular spacing (SD of 10%), respectively. It is observed from Fig. 2.5(a) that the proposed method has a slight advantage in precision over the entire range of simulated MSS values for a deterministic spacing. However, it is evident from Fig. 2.5(b) that as the spacing becomes irregular, our proposed method clearly outperforms the other methods since it shows a lower SD over the range of simulated MSS values and the variation of SD over the experimental range is also lower compared to the other methods.

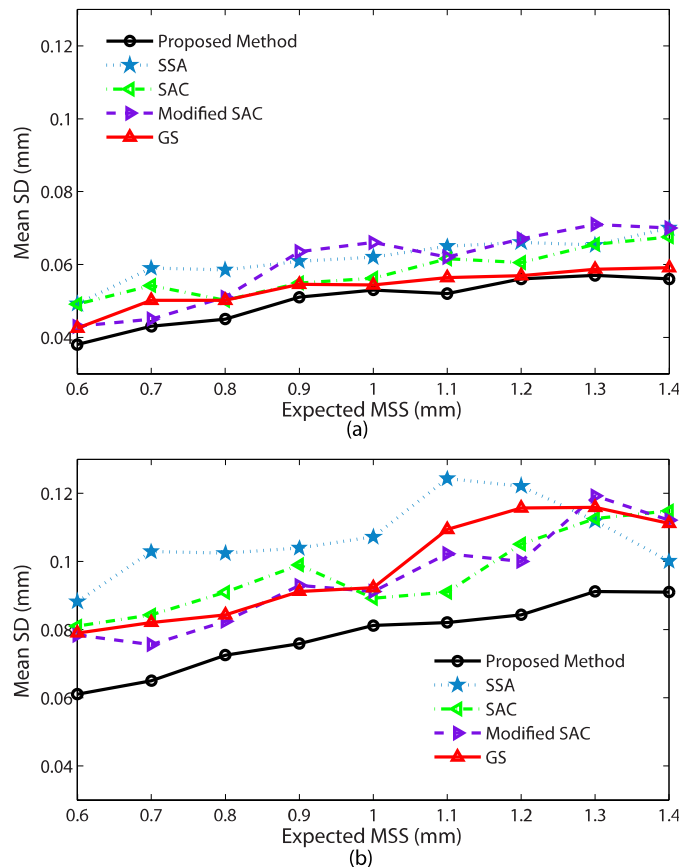


Figure 2.5: Variation of mean SD of the estimated MSS for different methods as the simulated MSS values are gradually increased for SD of regular scatterer spacing of (a) 0 % and (b) 10%.

The simulation results presented in Table 2.2 suggest that the conventional AR spectrum- and AR cepstrum-based techniques become less accurate as MSS values increase. It can be deduced from (2.10) that a higher MSS value results from peaks due to regular scatterers in the low frequency region of the spectrum. This points to the inability of the conventional AR spectrum- and AR cepstrum- based techniques in accurately detecting harmonics in the low frequency region of the spectrum, which can be attributed to the low frequency system effects. The conventional AR spectrum method makes no attempt to reduce the low frequency system effects. The conventional cepstrum-based technique, on the other hand, is known to reduce the low frequency system effects [4]. However, this method is based on oversimplifying assumptions which include an assumption that the PSFs and the tissue reflectivity functions (TRFs) reside in separate spectral bands and an assumption of minimum phase to reconstruct the PSF [65]. The cepstrum method also suffers from phase unwrapping problems [66]. The employed block based deconvolution technique, developed in our previous work [55], has been shown to solve many of these problems and hence, serves as a more effective deconvolution technique.

The results obtained from studying different simulation conditions as discussed in the FIELD II simulation setup section for the SSA, SAC, modified SAC, GS, and our proposed method are presented in Figs. 2.6(a), 2.6(b), and 2.6(c).

It is evident from Fig. 6(a) that although all methods perform well when the scattering amplitude of the regular scatterers is about half of that of the diffuse scatterers, the proposed method clearly outperforms the other methods when the relative reflectivity of the regular scatterers is much lower compared to the relative reflectivity of the diffuse scatterers. This can be attributed to the ability of EEMD in recovering the regular component even when it is corrupted by intense noise [16], [17].

Another important point to consider is the consistency of the spacing of the regular scatterers. In the results presented in Table 2.2, the MAPE values are calculated for regular scatterers with almost perfectly deterministic spacing. As can be seen from Fig. 2.6(b), the performance of all the methods along with our proposed method deteriorates as the scatterer spacing becomes more inconsistent. But the rise in MAPE

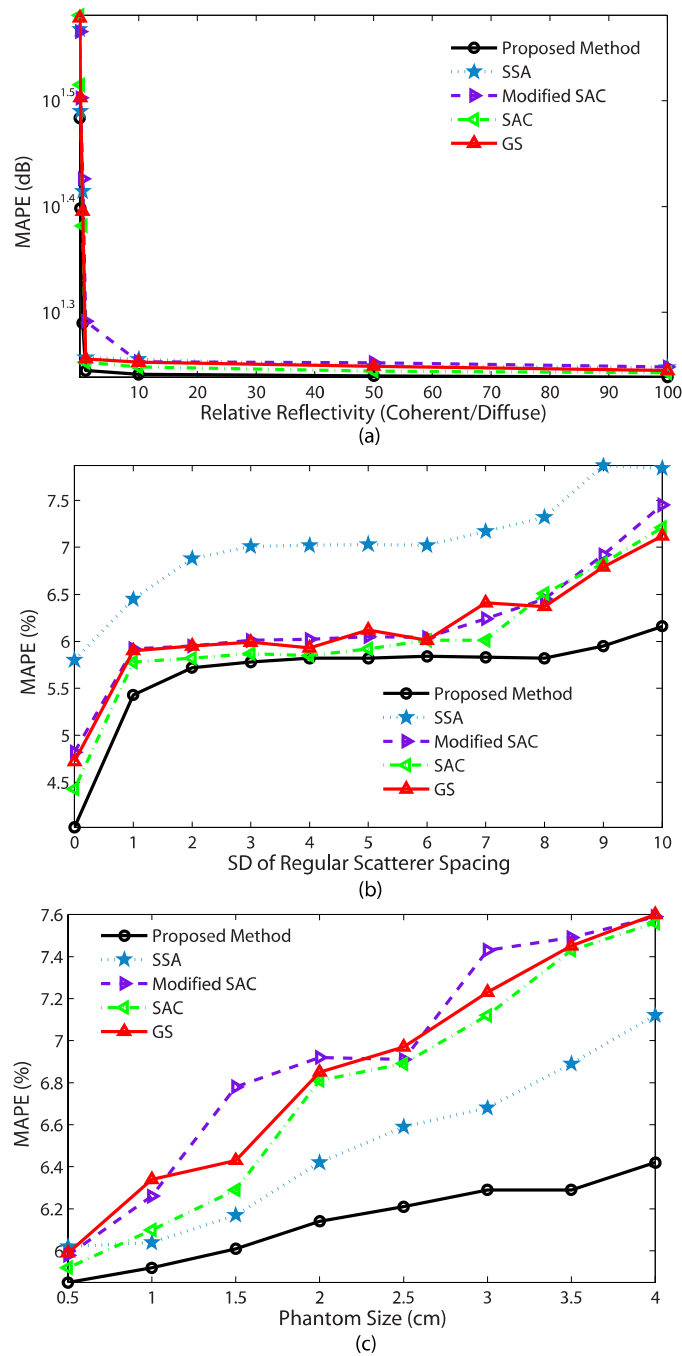


Figure 2.6: Effect of changing (a) the relative reflectivity of the regular and diffuse scatterers, (b) the consistency of spacing of the regular scatterers and (c) the phantom size for a simulated MSS value of 1 mm.

from 5.8% to about 6.16% for our proposed method is much lower compared to SSA, SAC, modified SAC, and GS. This can again be attributed to the ability of EEMD of recovering the regular component even when the periodicity of the component is weak [67].

Additionally, the axial size of the simulation phantom may effect the accuracy of the proposed method because of attenuation. As lesions may be located at different depths from the surface of the breasts, the attenuation effect may become significant with depth. The results, illustrated in Fig. 2.6(c), reveal that none of the proposed method, SSA, SAC, modified SAC or GS are adversely impacted by attenuation. However, the proposed method provides a clear advantage over the other methods. This can be attributed to the ability of EEMD in separating damped sinusoids [68].

2.6 Results on *in vivo* Breast Tissues

The MSS of *in vivo* breast data was estimated by our proposed algorithm as well as all the other techniques used in the preceding section for the simulated data. The results are summarized in Table 2.3.

The proposed method produces the MSS estimates with the lowest standard deviations for all types of tissue. Moreover, our proposed method, on our dataset, produces the maximum separation between the mean values of benign and malignant tissues. The range of MSS values obtained from *in vivo* data for normal, fibroadenoma and benign lesions, as shown in Table 2.3, is comparable to those reported by Bige *et al.* [4] and Taddayon *et al.* [5]. According to the best of our knowledge, the MSS values of inflammatory tissues and cyst have not been previously reported. The estimated MSS values for inflammatory tissues are slightly higher than those for normal tissues. The estimated MSS values of cysts are rather erratic (that is, the proposed technique often fails to produce any ESD estimates) and show a high standard deviation (SD) and hence, are not presented. This is true for all the techniques used to obtain the MSS estimates in Table 2.3. Therefore, it can be concluded that ESD estimates of cyst are of no diagnostic importance. This is in good concordance with the anatomy

Table 2.3: Estimated values of MSS with SD (in bracket) using different methods for different tissue types

Method	MSS (\pm SD) (mm)			
	Malignant	Fibroadenoma	Normal	Inflammatory
AR Spectrum [45]	0.97 (\pm 0.08)	1.00 (\pm 0.09)	1.03 (\pm 0.08)	1.01 (\pm 0.09)
AR Cepstrum [4]	0.87 (\pm 0.05)	0.91 (\pm 0.06)	0.93 (\pm 0.05)	0.92 (\pm 0.07)
SSA [14]	0.87 (\pm 0.05)	0.83 (\pm 0.05)	0.80 (\pm 0.05)	0.82 (\pm 0.06)
SAC [15]	0.65 (\pm 0.05)	0.67 (\pm 0.06)	0.70 (\pm 0.08)	0.69 (\pm 0.06)
Modified SAC [5]	0.75 (\pm 0.05)	0.77 (\pm 0.05)	0.80 (\pm 0.07)	0.78 (\pm 0.07)
GS [25]	0.70 (\pm 0.05)	0.72 (\pm 0.05)	0.74 (\pm 0.07)	0.74 (\pm 0.06)
Proposed Method	0.79 (\pm 0.04)	0.75 (\pm 0.03)	0.69 (\pm 0.03)	0.73 (\pm 0.04)

of the cysts as they are fluid-filled sacs and thus, scattering from cysts will largely be absent and some inconsistent scattering may occur due to debris (such as those present in complex cysts) [69]. The average MSS value for ROIs taken outside the lesions is estimated to be 0.70 (\pm 0.04) mm, corresponding to that of normal pathology.

Chapter 3

Effective Scatterer Diameter Estimation

The main focus of this Chapter is to discuss the formulation and theoretical aspects of the proposed effective scatterer diameter (ESD) estimation algorithm. The experiments performed on the tissue-mimicking phantoms (TMPs) and *in vivo* female breast data to estimate the ESD values using the proposed algorithm are discussed in detail here. The performance of this new estimator is analyzed by comparing its results with those obtained by the mean average squared deviation (MASD) based technique [10] and the conventional frequency domain method [9], as these methods also employ a Gaussian form factor to model tissue scattering like our proposed method.

3.1 Patient Data

The same data as used in the case of mean scatterer spacing (MSS) estimation, described in Chapter 2, has been used for ESD estimation as well.

3.2 TMP Data

A homogenous TMP, namely A, of Computerized Imaging Reference Systems Inc. (CIRS) of dimension $3 \times 4 \text{ cm}^2$ is used as the reference phantom. The TMP contains

glass beads as scatterers. The speed of sound in the TMP is around 1530 m/s. The TMP data was acquired using the same transducer settings as the *in vivo* data. For the estimation of average ESD we have used two homogeneous CIRS TMPs, namely A and B, which are inclusion-free, of dimensions $3 \times 4 \text{ cm}^2$ and one heterogeneous TMP, namely C, of dimensions $4.5 \times 4 \text{ cm}^2$, having an inclusion of diameter 1.4 cm. The actual average ESD, as supplied by the manufacturer, i.e, CIRS, are used as gold standards for performance evaluation of the ESD estimators. The ESD description of the experimental phantom datasets are presented in Table 3.1.

Table 3.1: Description of Experimental TMPs

TMP Dataset	Description	Average ESD (μm)	
		Inclusion	Background
A	Homogeneous	-	45
B	Homogeneous	-	45
C	Heterogenous	70	45

3.3 The ESD Estimation Technique

3.3.1 Preprocessing

A block diagram of our proposed ESD estimation algorithm is shown in Fig. 3.1. At first, 2-D regions of interest (ROI) are selected from suitable frames of the recorded RF data, in the same manner as discussed in the previous chapter for MSS estimation. For estimating the average ESD from TMP datasets A and B, 25 ROIs are selected, each of dimension $1 \times 1 \text{ cm}^2$, while for TMP dataset C, 10 ROIs are selected from outside the inclusion and 5 ROIs are selected within the inclusion. In TMP dataset C, we have also selected 5 heterogenous ROIs across the border of the inclusion such that the ROIs encompass both the inclusion and the background. As illustrated in

Fig. 3.1, deconvolution and filtering of the raw RF data serve as preprocessing steps. In this work, a multi-step system effect minimization scheme has been proposed with deconvolution and filtering serving as the first two steps. The last step of system effect minimization is normalization using a reference TMP as shown in the block diagram of Fig. 3.1. The RF data is deconvolved using the algorithm in [70] which removes the effect of the system PSF. After that, an ideal bandpass filter of frequency range 2 – 13 MHz is applied on the deconvolved data.

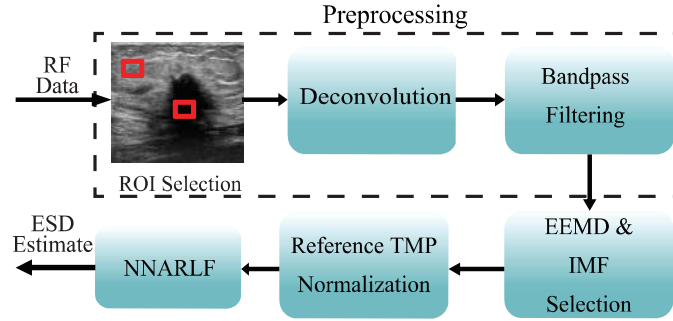


Figure 3.1: A block diagram illustrating our proposed algorithm. NNARLF refers to nearest neighborhood average regression line fitting

3.3.2 Proposed ESD Estimation Technique

The gated backscattered RF signal intensity, $W(f, z)$, at the transducer face can be expressed in the frequency domain as [71]

$$W(v) = T(f) \cdot D(f, z) \cdot A(f, z) \cdot S(f, D_{eff}, n_z), \quad (3.1)$$

where $T(f)$ represents the combined effect of the transmit pulse and the transducer sensitivity (electro-acoustic and acousto-electric transfer functions); $D(f, z)$ is the effect of diffraction; $A(f, z)$ is the cumulative attenuation in the soft tissue; $S(f, D_{eff}, n_z)$ represents the scattering properties of the tissue, including the effective scatterer diameter (D_{eff}), and acoustic concentration (n_z); z is the depth of the gated segment from the transducer face; and $v = \{f, D_{eff}, z, n_z\}$ is the set of variables on which the backscattered RF signal depends [13], [71]. We acquire RF signals from the tissue

sample and the reference TMP. The backscattered RF signal is then deconvolved and filtered. The resulting signal, in the frequency domain, is given by $W'(v)$. The cumulative attenuation, $A(f, z)$, in soft tissues is a function of frequency f and depth z , and can be expressed as [72]

$$A(f, z) = e^{-4\rho(f)z} = 10^{-2\rho(f)z/10}, \quad (3.2)$$

where $\rho(f)$ denotes the AC in unit Nepers/cm. It is reported in [72] that $\rho(f)$ demonstrates a linear frequency dependence. Therefore, it can be written as $\rho(f) = \beta \cdot f$, where β denotes the AC in Nepers/cm/MHz. By compensating for the effect of frequency dependent attenuation we get the compensated backscattered RF signal in the frequency domain as

$$\begin{aligned} W'_{comp}(v) &= W'(v)A_c(f, z) \\ &= T'(f) \cdot D'(f, z) \cdot S(f, D_{eff}), \end{aligned} \quad (3.3)$$

where $A_c(f, z)$ is the frequency dependent attenuation compensation function defined as

$$A_c(f, z) = A^{-1}(f, z); \quad (3.4)$$

and $T'(f)$ and $D'(f, z)$ are residual effects of the system PSF and diffraction, respectively, remaining after deconvolution and filtering. The AC and the average ESD of the TMP are known from the manufacturer's specifications. The method for estimating the AC of the tissues is described in [73]. To perform EEMD of the attenuation-compensated RF data, an ensemble of N_E random Gaussian noise, $g_p(n)$ ($p=1, \dots, N_E$) maintaining a signal-to-noise ratio of 30 dB with the attenuation-compensated signal in the time domain, $w'_{comp}(n)$, is added to $w'_{comp}(n)$. That is, the ensemble is given by [67]

$$w'_{pcomp}(n) = w'_{comp}(n) + g_p(n), p = 1, \dots, N_E. \quad (3.5)$$

After that, the EMD algorithm [16] is applied to each of the signals in the ensemble to decompose them into a sum of their IMFs, $c_{pj}(n)$, $j=1, \dots, K$, where K is the number of IMFs and a residue, $r_p(n)$, given by [67]

$$w'_{pcomp}(n) = \sum_{j=1}^K c_{pj}(n) + r_p(n), p = 1, \dots, N_E. \quad (3.6)$$

Finally, the IMFs using EEMD are obtained from the ensemble average [67]

$$\overline{c_j}(n) = \frac{1}{N_E} \sum_{p=1}^{N_E} c_{pj}(n), j = 1, \dots, K. \quad (3.7)$$

The IMFs responsible for diffuse scattering, $\overline{c_d}(n)$, $d=1, \dots, M$, where M is the number of IMFs (normalized) responsible for diffuse scattering, are then identified by the K-S test [62]. The IMFs are normalized using their amplitude. Normalization of the IMFs is done to ensure that no undue weight is given to any one of the IMFs. The attenuation-compensated RF data, for both the sample and the TMP, are then replaced by the summation of the IMFs responsible for diffuse scattering as

$$w' i_{comp}(n) = \sum_{d=1}^M \overline{c_d}(n), d = 1, \dots, M, \quad (3.8)$$

where, $w' i_{comp}(n)$ is the IMF-replaced signal.

Now, the attenuation compensated intensity spectra for the IMF replaced signals for the reference and sample tissues, in the frequency domain, can be rewritten as

$$W'_{R,comp}(v) = T'(f) \cdot D'_R(f, z) \cdot S_R(f, D_{eff,R}, n_{z,R}), \quad (3.9)$$

$$W'_{S,comp}(v) = T'(f) \cdot D'_S(f, z) \cdot S_S(f, D_{eff,S}, n_{z,S}), \quad (3.10)$$

where the subscripts R and S denote the reference and sample, respectively. For the same average sound speed in the reference and sample tissues, the diffraction terms can be considered as $D'_R(f, z) = D'_S(f, z)$. Finally, dividing (3.10) by (3.9), the normalized spectra, $W_n(v)$, is obtained as

$$W_n(v) = \frac{W'_{S,comp}(v)}{W'_{R,comp}(v)} = \frac{S_S(f, D_{eff,S}, n_{z,S})}{S_R(f, D_{eff,R}, n_{z,R})}. \quad (3.11)$$

Taking the logarithm on both sides of (3.11), we get

$$10 \log W_n(v) = 10 \log \frac{S_S(f, D_{eff,S}, n_{z,S})}{S_R(f, D_{eff,R}, n_{z,R})}. \quad (3.12)$$

A model for tissue scattering, $S(f, D_{eff}, n_z)$, in the frequency domain was developed in [74] for clinical array systems given by

$$S(f, D_{eff}, n_z) = \frac{185Lq^2(\frac{D_{eff}}{2})^6 n_z f^4}{[1 + 2.66(fq(\frac{D_{eff}}{2})^2)]} \times e^{-12.159f^2(\frac{D_{eff}}{2})^2} \quad (3.13)$$

with L the gate length (mm), q the ratio of aperture radius to distance from the region of interest, f the frequency in MHz, and D_{eff} the ESD in mm. The model was derived using a Gaussian form factor model and it has been previously established that a Gaussian form factor best models the scattering from human tissues [10], [43]. The quantity, n_z , is termed the acoustic concentration as defined in [10]. For sample and reference, the tissue scattering is then represented as

$$S_S(f, D_{eff,S}, n_{z,S}) = \frac{185Lq^2(\frac{D_{eff,S}}{2})^6 n_{z,S} f^4}{[1 + 2.66(fq(\frac{D_{eff,S}}{2})^2)]} \times e^{-12.159f^2(\frac{D_{eff,S}}{2})^2}, \quad (3.14)$$

$$S_R(f, D_{eff,R}, n_{z,R}) = \frac{185Lq^2(\frac{D_{eff,R}}{2})^6 n_{z,R} f^4}{[1 + 2.66(fq(\frac{D_{eff,R}}{2})^2)]} \times e^{-12.159f^2(\frac{D_{eff,R}}{2})^2}. \quad (3.15)$$

Dividing (3.14) by (3.15), and considering the fact that $2.66(fq\frac{D_{eff}}{2})^2 \ll 1$ [9], we get the normalized tissue scattering as

$$\frac{S_S(f, D_{eff,S}, n_{z,S})}{S_R(f, D_{eff,R}, n_{z,R})} = \frac{D_{eff,S}^6 n_{z,S}}{D_{eff,R}^6 n_{z,R}} e^{-3.03975(D_{eff,S}^2 - D_{eff,R}^2)f^2} \quad (3.16)$$

It is to be noted that for TMP datasets B and C, TMP dataset A has been used as reference while for TMP dataset A, TMP dataset B has been used as reference. Next, taking logarithm on both sides of (3.16) yields

$$10 \log \frac{S_S(f, D_{eff,S}, n_{z,S})}{S_R(f, D_{eff,R}, n_{z,R})} = 10 \log \frac{D_{eff,S}^6 n_{z,S}}{D_{eff,R}^6 n_{z,R}} - 13.20 \times (D_{eff,S}^2 - D_{eff,R}^2)f^2. \quad (3.17)$$

To estimate ESD, we fit a regression line through the usable (i.e., 6 dB) bandwidth of the normalized log scattering power spectrum. From (3.12), this is equivalent to fitting

a line through the usable bandwidth of the normalized log power spectrum. Assuming $f^2 = x$, the regression line can be expressed as

$$y = mx + c \quad (3.18)$$

with

$$y = 10 \log \frac{S_S(f, D_{eff,S}, n_{z,S})}{S_R(f, D_{eff,R}, n_{z,R})}, \quad (3.19)$$

$$m = -13.20 \times (D_{eff,S}^2 - D_{eff,R}^2), \quad (3.20)$$

$$c = 10 \log \frac{D_{eff,S}^6 n_{z,S}}{D_{eff,R}^6 n_{z,R}}. \quad (3.21)$$

Using average block power spectra generated from spatial signal blocks of sufficiently large size, ESD can be estimated from (3.20). On the other hand, probability that a single gated RF block includes regions of heterogenous tissues increases with the block size. Hence to trade-off between homogeneity in the spatial blocks and consistency of the estimated power spectrum, we use a weighted nearest neighbors algorithm. We assume that the ESD of the neighboring tissues of the scattering particles in the neighboring blocks are almost the same for their physical proximity. This is a more accurate representation of tissue structure and it is later established by generating the ESD map for a representative tissue. Therefore, to improve the reliability of the estimate, we calculate an average regression line as the weighted average of the regression lines of the neighboring blocks as

$$Y(i_s, j_s) = \frac{\sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} y(i_0, j_0) w^{(i_s, j_s)}(i_0, j_0)}{\sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} w^{(i_s, j_s)}(i_0, j_0)} \quad (3.22)$$

where $Y(i_s, j_s)$ denotes the weighted average value of $y(i_s, j_s)$, and $w^{(i_s, j_s)}(i_0, j_0)$ is the exponential weight function for an interrogated point (i_s, j_s) on the 2-D ESD map, defined as

$$w^{(i_s, j_s)}(i_0, j_0) = e^{-|\lambda_a(i_0-i_s)|-|\lambda_l(j_0-j_s)|}, \quad (3.23)$$

where λ_a and λ_l denote the weighting factors in the axial and lateral directions, respectively, and $i_s - L_a \leq i_0 \leq i_s + L_a$ and $j_s - L_l \leq j_0 \leq j_s + L_l$. L_a and L_l are the nearest

neighbors (NN) factors in the axial and lateral directions, respectively. From (3.23) it is evident that $w^{(i_s, j_s)}(i_0, j_0)$ has the maximum value (unity) at $(i_0, j_0) = (i_s, j_s)$. We define $w^{(i_s, j_s)}$ in a way such that in the averaging process, the logarithm of measured power spectrum of a neighboring window is properly weighted to have less contribution with increasing distance from the interrogated block. A 2-D weighted exponential neighborhood having $L_a = L_l = 5$ is illustrated in Fig. 3.2. The values on the weight axis are arbitrary but show an exponential decay as we move away from the interrogated window both axially and laterally.

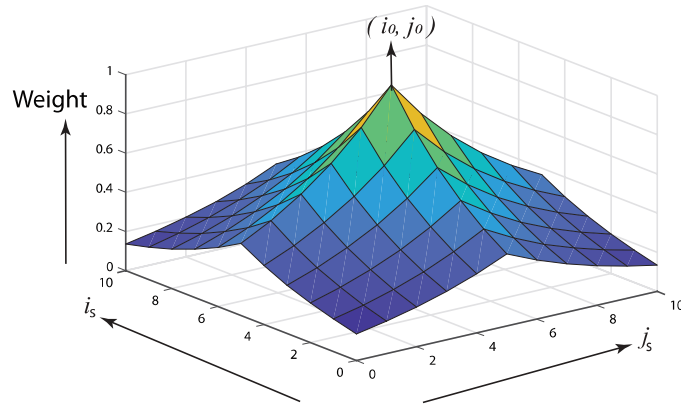


Figure 3.2: An illustration of the exponentially weighted neighborhood.

Substituting the value of y from (3.18) into (3.22), we get

$$Y(i_s, j_s) = \frac{A}{B}, \quad (3.24)$$

where

$$A = \sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} (m(i_0, j_0)x + c(i_0, j_0))w^{(i_s, j_s)}(i_0, j_0), \quad (3.25)$$

and

$$B = \sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} w^{(i_s, j_s)}(i_0, j_0). \quad (3.26)$$

If we define weighted average value of the slope as

$$M(i_s, j_s) = \frac{\sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} m(i_0, j_0)w^{(i_s, j_s)}(i_0, j_0)}{\sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} w^{(i_s, j_s)}(i_0, j_0)}, \quad (3.27)$$

and weighted average value of the intercept as

$$C(i_s, j_s) = \frac{\sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} c(i_0, j_0) w^{(i_s, j_s)}(i_0, j_0)}{\sum_{i_0=i_s-L_a}^{i_s+L_a} \sum_{j_0=j_s-L_l}^{j_s+L_l} w^{(i_s, j_s)}(i_0, j_0)}, \quad (3.28)$$

then (3.24) can be written in the form of a regression line given by

$$Y(i_s, j_s) = M(i_s, j_s)x + C(i_s, j_s). \quad (3.29)$$

Now, from the slope, M , of the regression line that fits (3.29) we can estimate the ESD (mm) at the point (i_s, j_s) using (3.20) as

$$D_{eff,S} = \sqrt{-\frac{M(i_s, j_s)}{13.20} + D_{eff,R}^2}. \quad (3.30)$$

In order to estimate the block power spectra of an interrogated block with higher resolution, the block is divided into 1-D segments with consecutive window segments in a block having 50% axial overlapping. The windowed segments are gated by the Hamming window. The block power spectra are calculated using the Welch method [75]. We select $L_a = L_l = 5$ as the NN factors to estimate the local ESD for a particular interrogated block. The impact of varying the NN factors on the ESD estimation is discussed later in this Chapter.

3.4 Experimental Phantom Results

To check the accuracy of our proposed ESD estimation technique, we use three CIRS experimental TMPs for which the average ESD values in the inclusion and background are available from the manufacturer. The average ESD of these phantoms are also estimated using the mean average square deviation (MASD) minimization based method [10], and the frequency domain method [9]. It is to be noted that a Faran form factor better models the scattering from glass beads present in the TMP [76]. But, since tissue scattering is more accurately modeled by a Gaussian form factor, to apply (3.11), a Gaussian form factor model has been applied for the TMPs as well.

Moreover, Gaussian form factor models have also been previously used to model the scattering from TMPs [77]. Table 3.2 presents the actual average ESD and the average ESD estimated using our proposed algorithm as well as the other techniques for the experimental phantoms. The average ESD represents the average of the ESD values estimated from each of the ROIs within a particular TMP. It is evident that our proposed algorithm estimates the ESD for all three TMP data with a higher degree of accuracy compared to the other methods as reflected by a lower mean absolute percentage error (MAPE) value. Moreover, our method also shows a lower standard deviation (SD) of estimates compared to the other methods.

Table 3.2: Estimated ESD values (in μm) with SD (in bracket) from experimental TMPs

Method	TMP Dataset	Average ESD of Background (μm) ($\pm\text{SD}$)	MAPE (%)	Average ESD inside Inclusion (μm) ($\pm\text{SD}$)	MAPE (%)
MASN Based Method [10]	A	54.82(± 8.88)	21.82	-	-
	R	52.18(± 8.09)	15.96	-	-
	C	50.91(± 7.99)	13.13	74.91(± 7.87)	7.01
Frequency Domain Method [9]	A	53.84(± 7.11)	19.64	-	-
	R	50.10(± 7.07)	11.33	-	-
	C	48.97(± 7.39)	8.82	73.61(± 7.09)	6.71
Proposed Method	A	47.02(± 5.89)	4.49	-	-
	R	47.76(± 6.01)	6.13	-	-
	C	47.78(± 5.85)	6.18	73.02(± 6.33)	5.15

We have also estimated the ESD from the heterogenous ROIs of TMP C using our proposed method. A scatter plot showing how the ESD varies as we move laterally from a region just outside the inclusion across the border of the inclusion to a region within the inclusion is shown in Fig. 3.3. It is evident from the figure that our proposed method is able to reliably estimate the ESD in the homogenous regions inside and outside the inclusion with a mean value of approximately 48 μm outside the inclusion (represented by a red line) and a mean value of approximately 74 μm within the inclusion (represented by a green line). There is a sharp change in the average ESD values across the border of the inclusion, as expected. Now, in order to study how the

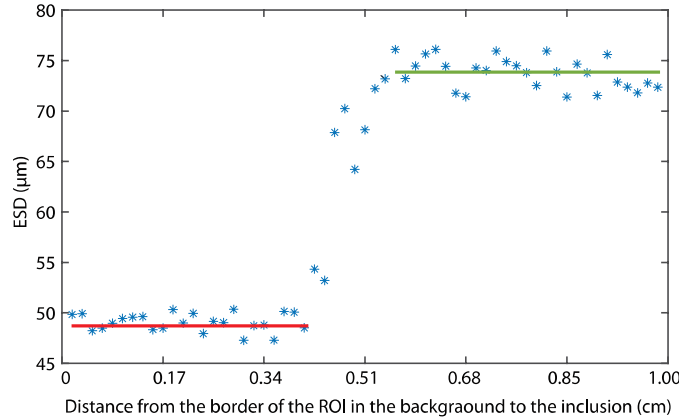


Figure 3.3: Scatter plot showing the variation of ESD across different windows on moving laterally from the edge of the ROI in the TMP background to the edge of the ROI within the inclusion.

different steps in the proposed algorithm impacts the accuracy of the ESD estimates, the ESD was estimated for several ROIs of TMP datasets A, B, and C by removing the different steps shown in the flow chart of Fig. 3.1 one by one while retaining the others. The results are presented in Fig. 3.4 in the form of a bar plot. It can be seen that the overall estimation accuracy is most adversely impacted on removing the EEMD step. There is also a slight rise in the SD of estimates on removing the EEMD step. Removal of the any one of the system effect minimization steps, i.e, deconvolution, filtering, and normalization using a reference TMP, also noticeably impacts the ESD estimation accuracy, with a slight increase in the SD of ESD estimates in each case. The weighted neighborhood step seems to have the least impact on the overall accuracy. However, the SD of the estimates is seen to rise more significantly from the proposed method. The first column of the bar plot presents the performance of the proposed algorithm on the same datasets for comparison. It can be inferred from the plot that the EEMD step has the most impact on the accuracy of the ESD estimation. Moreover, it also justifies the use of a multi-step system effect minimization technique since removal of any one of these steps has detrimental effect on the overall ESD estimation accuracy.

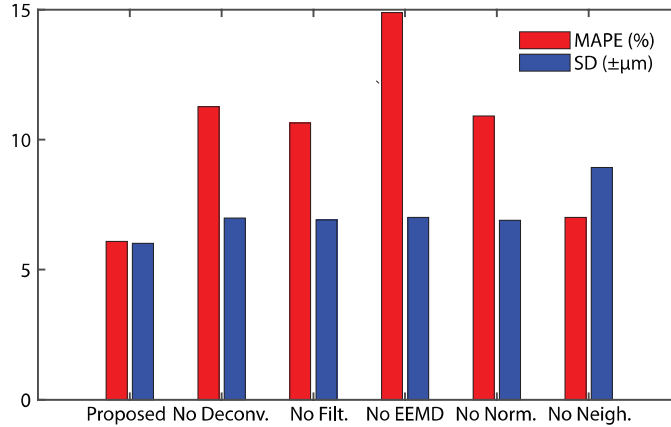


Figure 3.4: Bar plot showing the impact on the percentage error and SD of the ESD estimates from the experimental TMPs on removing the different steps from our proposed algorithm.

3.5 Results on *in vivo* Breast Tissues

The results of estimating the ESD of *in vivo* breast tissues using various techniques are presented in Table 3.3. The ESD values estimated using our proposed algorithm for fibroadenoma and malignant breast tissues fall within the range of ESD values previously reported in the literature [5]. Similar to MSS, the ESD values of inflammatory lesions and cystic lesions have not been previously reported. The ESD values in the region outside the lesions are consistent in the range between 70 – 80 μm , which is similar to normal tissues. The ESD estimates for inflammatory tissues show little or no deviation from this range. The estimated ESD values of cysts are again erratic, as is the case for MSS, and hence, are not presented. This is true for all the techniques used to obtain the ESD estimates in Table 3.3.

Furthermore, it is seen that the average ESD value of malignant lesions is greater than that of benign lesions which also conform with the previously reported results [5]. The estimated average ESD values using the MASD-based method [10] and the original frequency domain method proposed in [9] are presented in the first and second rows of Table 3.3. We see that the SD of ESD estimates are significantly higher than that of our proposed method and the separation between the average ESD values for

Table 3.3: Estimated values of ESD with SD (in bracket) using different methods for different tissue types

ESD (\pm SD) (μm)							
	Malignant		Fibroadenoma		Adipose		Normal
Method	in	out	in	out	in	out	
MASTN based method [10]	100.67 (\pm 21.13)	86.12 (\pm 8.41)	94.09 (\pm 18.17)	82.10 (\pm 7.24)	88.31 (\pm 11.34)	90.35 (\pm 10.01)	76.13 (\pm 7.01)
Frequency domain method [9]	109.21 (\pm 17.34)	80.04 (\pm 8.21)	101.41 (\pm 13.12)	79.31 (\pm 8.54)	81.24 (\pm 9.81)	82.45 (\pm 11.01)	75.88 (\pm 6.74)
Proposed Method	123.05 (\pm 8.85)	74.90 (\pm 4.19)	98.71 (\pm 9.55)	74.89 (\pm 4.11)	75.72 (\pm 4.09)	75.77 (\pm 4.07)	75.12 (\pm 4.01)

malignant lesions and fibroadenomas are also smaller.

In order to show the correlation between the ESD estimates and the histology images, the representative hematoxylin and eosin (H & E) histology sections of fibroadenomas, malignant tissues and inflammatory tissues are presented in Figs. 3.5(a)–(c), respectively. In the histology images, the microlobules are purple stained while the stroma are pink stained and the fat is white. It is evident that the size of the micro-lobules in the malignant tissue section are larger. The sizes of the lobules in the malignant tissue section also vary more widely than in the fibroadenoma section. The section for inflammatory tissues is typical of what is expected for normal tissues. On measuring the average sizes of the glandular structures in the fibroadenoma and malignant sections, with the help of a microscope, they are found to be $95\pm 17 \mu\text{m}$ and $149\pm 24 \mu\text{m}$, respectively. In comparison, using our proposed method, the average ESDs corresponding to fibroadenomas and malignant tissues are found to be $98.71\pm 9.55 \mu\text{m}$ and $123.05\pm 8.85 \mu\text{m}$, respectively. Hence, the values obtained from the sole histology image of each type are in fairly good concordance with the values estimated from our proposed method.

An important factor that has to be taken into consideration while estimating ESD from *in vivo* tissues is the impact of the kernel size of the weighted exponential neigh-

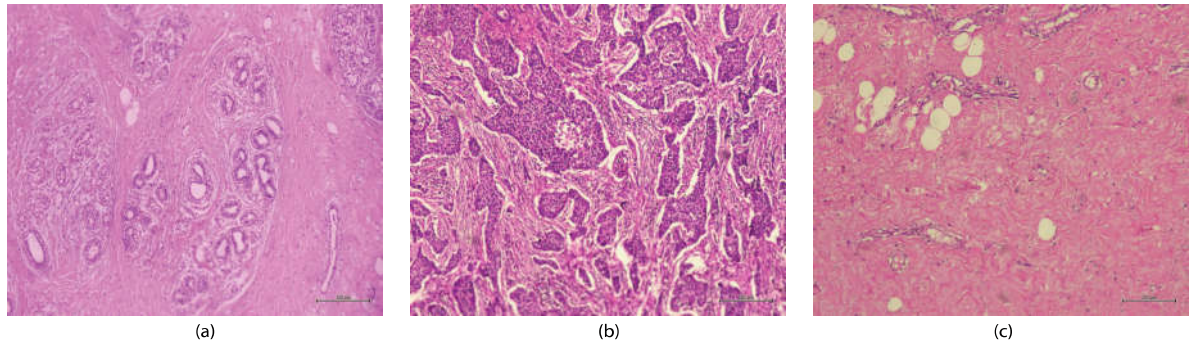


Figure 3.5: Histology images of (a) fibroadenoma, (b) malignant and (c) inflammatory breast tissue. Scale bar: $100 \mu\text{m}$ in (a) and $200 \mu\text{m}$ in (b) and (c).

neighborhood. The algorithm involves the use of a 5×5 weighted exponential neighborhood. The use of such a neighborhood allows the modeling of tissue homogeneity over a small region rather than a large spatial block where the tissue becomes more heterogeneous. To investigate the impact of the size of the neighborhood on the ESD estimates, the ESD is estimated using our proposed algorithm for normal tissues for no neighborhood, a neighborhood of size 3×3 , and a neighborhood of size 8×8 . It has already been stated that the original results are produced for a neighborhood of size 5×5 . The results are shown in Table 3.4.

Table 3.4: Estimated values of ESD with SD (in bracket) for normal tissues with different neighborhood sizes.

Size of Neighborhood	ESD (\pm SD) (μm)
No Neighborhood	76.89 (\pm 8.56)
3×3	74.99 (\pm 4.09)
5×5	75.12 (\pm 4.01)
8×8	76.46 (\pm 7.56)

It is clear from the table that the choice of neighborhood size does not greatly effect the average value of the ESD estimates. However, the SD of estimates seem to decrease significantly for a neighborhood size close to our selected one. A large neighborhood

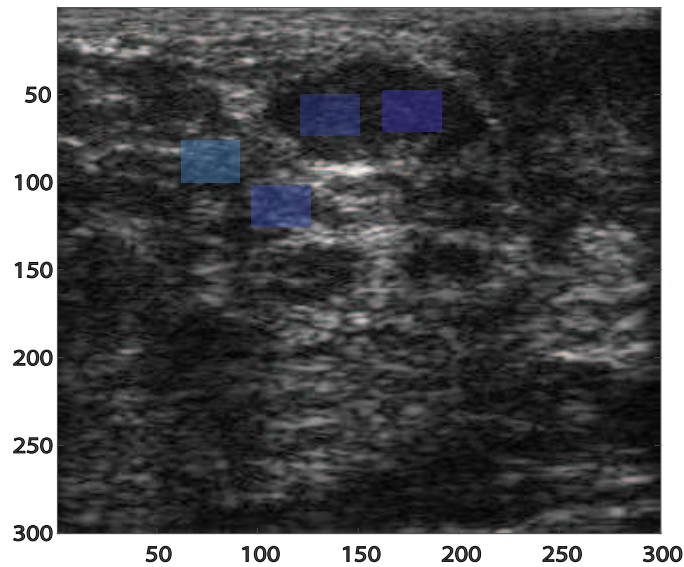


Figure 3.6: ESD map for four ROIs for a representative fibroadenoma tissue. A deeper shade of blue indicates higher ESD values. The values on the axes are arbitrary.

size or no neighborhood increases the SD of estimates and this could be attributed to the increased heterogeneity for a large spatial block of tissue. This observation is consistent to that obtained for experimental TMPs where removal of the weighted neighborhood step adversely impacted the SD of estimates. To further substantiate our argument for choosing an exponentially weighted neighborhood, we produce, in Fig. 3.6, an ESD map for a representative fibroadenoma tissue for 2 ROIs taken within the border of the lesion and 2 ROIs taken outside the lesion. It is clear that the ESD values are consistent (having a low SD) across the ROIs (both inside and outside the lesion) and hence, a weighted exponential neighborhood is a good model of tissue structure.

Chapter 4

Breast Lesion Classification

In this Chapter, we discuss the binary (benign-malignant) classification performance of our proposed ultrasonic micro-parameters, mean scatterer spacing (MSS) and effective scatterer diameter (ESD). We then combine it with 27 previously reported ultrasound bi-modal macro-parameters, estimated from the ultrasound B-mode (UB) images and ultrasound elastography (UE), to form a unique hybrid micro-macro feature set, and obtain the classification results. The classification performance is evaluated using a wrapper based scheme, a genetic algorithm (GA) based optimization and also, an empirical mode decomposition-discrete wavelet transform (EMD-DWT) based scheme.

4.1 Classification Dataset

Out of the 245 RF data records, 44 normal patients and 42 cystic lesions are excluded since normal data can be classified directly from the UB image from the absence of a growth and micro-parameters cannot be reliably estimated for cystic lesions. The total dataset for classification, therefore, consists of 159 RF data, which include 56 malignant lesions, 79 fibroadenomas, and 24 inflammatory lesions.

4.2 Micro- and Macro-Parameter Based Classification of Breast Lesions

In [18], bi-modal (UB and UE) macro-parameters have been successfully used for classification of breast lesions. In this thesis, we attempt to further improve the classification accuracy achieved in [18] by extending the feature space with the micro-parameters extracted here to produce a unique hybrid feature set of micro- and macro-parameters. Bi-modal QUS based macro-parameters have been previously reported in the literature [78]–[91]. The macro-parameters are mainly composed of shape, strain, boundary, margin, contrast, and echo-pattern based features and are estimated from the UB images and from UE. Out of the full-set of macro-parameters, 27 features that have shown promise in breast lesion classification are listed in Table 4.1. In this thesis, these features are combined with MSS and ESD for binary classification of breast lesions.

4.3 Classification Techniques

The classification performances of MSS and ESD alone, and the combination of MSS and ESD (both normalized) are evaluated with the help of support vector machine (SVM), K-nearest neighbor (KNN), linear discriminant analysis (LDA), multinomial logistic regression (MNR), and Naïve Bayes (NB) classifiers. The best classification performance obtained has been reported in this work. Here, we used the commonly employed “one-versus-all” or OVA based classification technique [92]. The total dataset of MSS/ESD values are subdivided into 5 groups, each group containing a mixture of benign and malignant data. One of the 5 groups is used to train a statistical classifier and this classifier is used to test (classify) the other 4 groups. In this way, the classifier is trained by each of the 5 groups in turn while the other 4 groups are tested.

The normalized hybrid feature set is first passed through a wrapper-based feature reduction scheme and classified by using a K-fold cross-validation technique described in [92], with the help of SVM, KNN, LDA, MNR, and NB classifiers. Again, the best classification performance obtained has been reported.

Table 4.1: List of bi-modal QUS macro-paramters

Feature Type	Real-valued feature
	Edge diffusivity (ED) [78]
Boundary contrast Descriptor	Lesion boundary feature (LB_d) [79]
	Normalized Radial Gradient (NRC) [80]
	Margin definition (M_{defn}) [81]
Lesion contrast	Co-contrast (C_{cnrst}) [82]
Posterior acoustic feature	Posterior acoustic feature (PS_d) [79]
Texture descriptor	Four neighborhood pixel algorithm (FNPA) [83]
Roughness descriptor	Hurst coefficient (H_{coeff}) [84]
Tissue heterogeneity	Gray-level non-zero heterogeneity (σ_{graynz}) [85]
Tissue echogeneity	Gray-level non-zero echogeneity (μ_{graynz}) [85]
	Morphometric features
	Shape asymmetry factor (SAF) [78]
	Aspect ratio (ASR) [86]
	SD of normalized radial length (σ_d) [87]
	Roundness ($R_{rndness}$) [88]
Shape Descriptor	Compactness (C_{mpct}) [81]
	Convexity (C_{nvx}) [89]
	Solidity (S_{ldt}) [88]
	Form factor (F_{factor}) [88]
	Shape factor (S_{PR}) [79]
	Margin feature (M_{UA}) [79]
Contour Descriptor	Circularity (C) [90]
	Mean of normalized radial length (d_{mean}) [87]
	Outside area ratio (A_{out}) [87]
	Semi-minor axis length of fitted ellipse(S_b) [79]
	Orientation of fitted ellipse(O_E) [79]
Ultrasonic Elastography Feature	Strain ratio (SR) [91]
	Area Ratio (AR) [91]

Moreover, a GA based optimization has also been employed for classification. The technique was previously proposed in [78]. Here, 11 out of the 29 parameters that demonstrate the best classification performance independently and also combined are chosen. The individual and combined performances of the micro- and macro-parameters are evaluated using the linear classifiers mentioned previously. The 11 parameters include 10 macro-parameters, namely, SR , ED , SAF , ASR , C_{nvx} , M_{UA} , AR , M_{defn} , and S_{ldt} , S_{PR} , and 1 micro-parameter, namely, ESD. A modified hybrid bimodal multiparameter

characterization index, CI , is defined as

$$\begin{aligned}
 CI = & \frac{1}{\sum_{i=1}^{11} w_i} [w_1.SR + w_2.ED + w_3.SAF \\
 & + w_4.ASR + w_5.C_{nvx} + w_6.M_{UA} \\
 & + w_7.AR + w_8.M_{defn} + w_9.S_{ldt} + w_{10}.S_{PR} + w_{11}.ESD], \\
 & 0 \leq w_i \leq 1. \text{ for all } i
 \end{aligned} \tag{4.1}$$

The weight ($\mathbf{w} = [w_1, w_2, \dots, w_n]^T$) values are selected using a GA based optimization algorithm. The weights are selected based on the maximization of the cost function, J_1 , defined as

$$J_1 = \text{sensitivity} + \text{specificity} + \text{accuracy}, \tag{4.2}$$

with the help of the GA as

$$\mathbf{w}_{\text{ga}} = \arg \max_{0 \leq w_i \leq 1} (J_1). \tag{4.3}$$

For all indices employing the GA, a mutation rate of 0.02 and a crossover rate of 0.8 are used. The classification is carried out again with the help of an OVA based technique as described in [92].

In addition, an EMD-DWT based scheme has also been used to evaluate the classification performance. This technique has been previously proposed in [18], which utilizes an EMD and DWT based feature transform method followed by a wrapper-based subset selection scheme. However, [18] used only the macro-parameters for classification of breast lesions. We ameliorate this technique using our hybrid feature set. It has been established in [18] that the transformed feature set using EMD and DWT is dependent on the order of the original features. Hence, at first, in order to select the best transform domain features, 2000 random sequences of the original features are produced and the performance of the reduced transform domain feature set (reduced using wrapper) are measured in terms of Sum_5 (the sum of sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV)). This single quality metric, Sum_5 , is calculated with a 10-fold cross validation. The order of the original feature set at which the transform domain feature set attains the maximum

Sum_5 , is selected.

The classification is evaluated using a number of performance metrics for each of the above mentioned techniques. The classification results include true positive (TP), true negative (TN), false positive (FP), false negative (FN), sensitivity (Sens.), specificity (Spec.), accuracy (Acc.), positive predictive value (PPV), negative predictive value (NPV), the sum of the last five parameters, Sum_5 , and Matthew's correlation coefficient (MCC) as quality metrics. TP refers to a diagnostic modality correctly identifying a malignant lesion, TN refers to a diagnostic modality correctly identifying a benign lesion, FP refers to a diagnostic modality incorrectly characterizing a benign lesion as a malignant lesion, while FN refers to a diagnostic modality incorrectly characterizing a malignant lesion as a benign lesion. From TP, TN, FP, and FN, sensitivity, specificity, accuracy, PPV, NPV and MCC can be computed as [93], [94], [95],

$$Sensitivity = \frac{TP}{TP + FN} \quad (4.4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (4.5)$$

$$Accuracy = \frac{TP + TN}{TN + FN + TP + FP} \quad (4.6)$$

$$PPV = \frac{TP}{TP + FP} \quad (4.7)$$

$$NPV = \frac{TN}{TN + FN} \quad (4.8)$$

$$MCC = \frac{(TP \times TN - FP \times FN)}{[(TP + FP)(TP + FN)(TN + FP)(TN + FN)]^{\frac{1}{2}}}. \quad (4.9)$$

Higher values of sensitivity, specificity, accuracy, PPV, and NPV indicate better classification performance. For MCC, +1 indicates a perfect prediction, 0 indicates a uniform random prediction, and -1 indicates an inverse prediction [94].

4.4 Classification Results

The classification results are presented in Table 4.2. The first three rows show the classification results obtained using only MSS, only ESD, and the combination of MSS and

Table 4.2: Breast lesion classification results obtained for the different micro- and macro-parameters

Features	TD	TN	FP	FN	Sens. (%)	Spec. (%)	Acc. (%)	DDV (%)	NDV (%)	Sum ₅	MCC
MSS	42	72	31	14	75	69.90	71.70	57.53	83.72	357.86	0.4304
ESD	51	99	4	5	91.07	96.12	94.34	92.73	95.19	469.44	0.8755
MSS and ESD	54	98	5	2	96.43	95.15	95.60	91.53	98.00	476.69	0.9054
Hybrid (wrapper)	54	101	2	2	96.43	98.06	97.48	96.43	98.05	486.46	0.9449
Hybrid (GA)	55	100	3	1	98.21	97.08	97.48	94.83	99.01	486.62	0.9457
Hybrid (EMD-DWT)	55	101	2	1	98.21	98.06	98.11	96.49	99.02	489.90	0.9589

ESD, respectively. The next three rows show the classification results on the proposed hybrid feature set using the three techniques discussed previously. A classification scheme based on ESD alone produces sensitivity, specificity, and accuracy values of 91.07%, 96.12%, and 94.34%, respectively. The classification performance of MSS on this dataset is not entirely satisfactory. But, on combining ESD with MSS, we obtain improved sensitivity, specificity, accuracy, and MCC values of 96.43%, 95.15%, and 95.60%, and 0.9054, respectively. As discussed before, the classification performance of the micro-parameters (ESD and MSS) have been evaluated directly using SVM, LDA, MNR, KNN, and NB classifiers. In Table 4.2, the reported Sum_5 value of 476.69 is obtained for a LDA classifier and represents the best Sum_5 value obtained out of all the above mentioned classifiers. The average Sum_5 value (\pm SD), found by averaging the Sum_5 values obtained from each of the mentioned classifiers, is 473.02 (\pm 3.47). This indicates that the classification performance obtained using the micro-parameters is fairly stable. It is evidently clear from the fourth row of the table that using a hybrid feature set significantly improves the obtained result. The wrapper based scheme increases the sensitivity, specificity, and accuracy values to 98.06%, 97.48%, and 96.43%, respectively. This result is obtained using a 9-fold backward elimination feature reduction scheme using a LDA classifier with an average of 13 features selected. This represents the best result obtained over different folds, different classifiers, and the two schemes for wrapper based feature reduction (forward/backward). It is seen from the fifth row of the table that the results improve slightly on using the GA optimization scheme. Furthermore, on using the EMD-DWT scheme, the Sum_5 value increases to

489.90. The obtained classification result, using the EMD-DWT scheme, is the best reported result, till date, for any QUS micro/macro parameter based breast lesion classification scheme. The use of Sum_5 to indicate the classification performance has been previously established in [18], [60], and [78]. In this work, we have also used MCC to evaluate the classification performance which is shown to have a high degree of correlation with the Sum_5 values. That is, a higher Sum_5 value generally leads to a MCC value closer to +1.

Chapter 5

Conclusion, Limitation and Future Scope

5.1 Conclusion

This thesis has presented novel methods for estimating the mean scatterer spacing (MSS) and effective scatterer diameter (ESD) by separation of the coherent and diffuse component of backscattered data using ensemble empirical mode decomposition (EEMD). Breast lesion classification using a novel hybrid of micro- and macro-parameters feature set has also been proposed. The proposed MSS and ESD estimation techniques have produced reliable MSS and ESD estimates as exhibited by accurate estimation of MSS from simulation phantoms and ESD from experimental TMPs. These methods are shown to be more superior compared to the existing techniques for MSS and ESD estimation, as reflected by lower values of mean absolute percentage error (MAPE) on estimating MSS values from simulation phantoms and ESD values from experimental phantoms. Furthermore, when the MSS is fused and 27 other macro-parameters estimated from the ultrasound bi-modal (UB) images and ultrasound elastography (UE) for binary classification of 159 breast lesions, very high values of sensitivity, specificity, accuracy, and MCC are obtained. Therefore, our proposed MSS and ESD estimation technique, along with the classification based on our hybrid micro- and macro-

parameter based feature set, show promises to be used as computer-aided diagnosis (CAD) tools for breast lesion classification.

5.2 Limitation and Future Scope

The main concentration of this research work is set on the accurate determination of MSS and ESD as clinical diagnostic parameters. However, other micro-parameters like the effective acoustic concentration (EAC), and the attenuation coefficient (AC), need to be explored, to further extend the feature space. Additionally, the theoretical power spectrum used for ESD estimation employs a Gaussian form factor to model tissue backscattering. While the Gaussian form factor is a fair representation of scattering from tissue structure, other adequate models like the Faran form factor and the fluid-filled spheres exist, and theoretical power spectrums need to be developed from these models to adequately compare the performance of different models for ESD estimation. In addition, preliminary histopathology studies have been carried out in this work to correlate the ESD estimates with those obtained from microscopy. However, the true gold standard for micro-parameter estimation must be based on extensive microscopy analysis, which could not be carried out in this work because of labeled histopathological slides not being available for the patients in the dataset. Furthermore, a drawback of the micro-parameters used in this paper is that they produce erratic values for cysts. This erratic behaviour has already been explained through the idea that cysts are fluid-filled sacs which do not have a consistent scattering signature. Thus, cysts have been excluded while obtaining the classification results in this work. As cysts are classified as benign lesions, a micro-parameter based classification scheme will be, therefore, unable to characterize cysts correctly. It is to be noted, however, that the macro-parameters can be accurately estimated for cysts since the macro-parameters are morphological features derived from the UB images and UE. In fact, macro-parameters produce a better classification performance if cysts are included. A macro-parameter based classification scheme on the same dataset (with cysts included) produced a Sum_5 value of 476.50 [78]. Also, some strain imaging techniques have been developed which can

successfully characterize cysts [63], [96], [97] and hence, may be used in conjunction with the micro-parameters for breast lesion classification. Moreover, the classification results obtained using the wrapper based schemes employ a number of parameters. The presented results represent the best results obtained over different folds, different classifiers, and the two schemes for wrapper based feature reduction (forward/backward). The results are seen to deteriorate on changing these conditions. Future works should, therefore, look to reduce the sensitivity of the classification performance to these parameters. Lastly, current trends point to a deep learning based approach for ultrasonic breast lesion classification [98]. While the dataset used in this work is still too limited for adopting such an approach, the proposed hybrid feature set may have potential to be used in conjunction with deep neural networks to develop a more robust CAD tool for breast lesion classification, if a larger dataset becomes available.

List of Publications

Journal

1. **N. I. Nizam**, S. K. Alam, and M. K. Hasan, “EEMD Domain AR Spectral Method for Mean Scatterer Spacing Estimation of Breast Tumors From Ultrasound Backscattered RF Data”, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 64, no. 10, pp. 1487–1500, 2017.

International Conference (1-page extended abstract)

1. **N. I. Nizam**, S. R. Ara, and M. K. Hasan, “Ultrasound Micro-Macro Parameter Based Breast Lesion Classification using GA Optimization and EMD-DWT”, Accepted, *IEEE-EMBS Conference on Biomedical and Health Informatics*, Illinois, Chicago, May 19–May 22, 2019.

Submitted Journal

1. **N. I. Nizam**, S. R. Ara, and M. K. Hasan, “Ultrasonic Estimation of Effective Scatterer Diameter for Micro-Macro Parameter Based Improved Breast Lesion Classification,” *IEEE Journal of Biomedical and Health Informatics*, 2019.

Bibliography

- [1] R. L. Siegel, K. D. Miller, and A. Jemal, “Cancer statistics, 2019,” *CA: a cancer journal for clinicians*, vol. 69, no. 1, pp. 7–34, 2019.
- [2] M. Calnan, “The health belief model and participation in programmes for the early detection of breast cancer: a comparative analysis,” *Social Science & Medicine*, vol. 19, no. 8, pp. 823–830, 1984.
- [3] J. C. Bloodgood, “Biopsy in breast lesions: In relation to diagnosis, treatment and prognosis,” *Annals of surgery*, vol. 102, no. 2, p. 239, 1935.
- [4] Y. Bige, Z. Hanfeng, and W. Rong, “Analysis of microstructural alterations of normal and pathological breast tissue in vivo using the ar cepstrum,” *Ultrasonics*, vol. 44, no. 2, pp. 211–215, 2006.
- [5] H. Tadayyon, A. Sadeghi-Naini, L. Wirtzfeld, F. C. Wright, and G. Czarnota, “Quantitative ultrasound characterization of locally advanced breast cancer by estimation of its scatterer properties,” *Med. Phys.*, vol. 41, no. 1, 2014.
- [6] K. Nam, J. A. Zagzebski, and T. J. Hall, “Quantitative assessment of in vivo breast masses using ultrasound attenuation and backscatter,” *Ultrason. Imaging*, vol. 35, no. 2, pp. 146–161, 2013.
- [7] F. L. Lizzi, M. Ostromogilsky, E. J. Feleppa, M. C. Rorke, and M. M. Yaremko, “Relationship of ultrasonic spectral parameters to features of tissue microstructure,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 34, no. 3, pp. 319–329, 1987.

- [8] W. K. Moon, C.-M. Lo, J. M. Chang, C.-S. Huang, J.-H. Chen, and R.-F. Chang, "Quantitative ultrasound analysis for classification of bi-rads category 3 breast masses," *Journal of Digital Imaging*, vol. 26, no. 6, pp. 1091–1098, 2013.
- [9] M. L. Oelze, J. F. Zachary, and W. D. O'Brien Jr, "Characterization of tissue microstructure using ultrasonic backscatter: Theory and technique for optimization using a gaussian form factor," *J. Acoust. Soc. Am.*, vol. 112, no. 3, pp. 1202–1211, 2002.
- [10] M. L. Oelze and W. D. O'Brien Jr., "Method of improved scatterer size estimation and application to parametric imaging using ultrasound," *J. Acoust. Soc. Am.*, vol. 112, no. 6, pp. 3053–3063, 2002.
- [11] G. Georgiou, F. S. Cohen, C. W. Piccoli, F. Forsberg, and B. B. Goldberg, "Tissue characterization using the continuous wavelet transform part II: application on breast rf data," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 48, no. 2, pp. 364–373, 2001.
- [12] K. A. Wear, R. F. Wagner, M. F. Insana, and T. J. Hall, "Application of autoregressive spectral analysis to cepstral estimation of mean scatterer spacing," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 40, no. 1, pp. 50–58, 1993.
- [13] Y. Labyed and T. A. Bigelow, "A theoretical comparison of attenuation measurement techniques from backscattered ultrasound echoes," *J. Acoust. Soc. Am.*, vol. 129, no. 4, pp. 2316–2324, 2011.
- [14] W. C. D. A. Pereira and C. D. Maciel, "Performance of ultrasound echo decomposition using singular spectrum analysis," *Ultrasound in Med. & Biol.*, vol. 27, no. 9, pp. 1231 – 1238, 2001.
- [15] T. Varghese and K. D. Donohue, "Estimating mean scatterer spacing with the frequency-smoothed spectral autocorrelation function," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 42, no. 3, pp. 451–463, 1995.

- [16] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.
- [17] N. E. Huang and Z. Wu, "A review on hilbert-huang transform: Method and its applications to geophysical studies," *Rev. Geophys.*, vol. 46, no. 2, p. 2007RG000228, 2008.
- [18] S. R. Ara, S. K. Bashar, F. Alam, and M. K. Hasan, "EMD-DWT based transform domain feature reduction approach for quantitative multi-class classification of breast lesions," *Ultrasonics*, vol. 80, pp. 22–33, 2017.
- [19] L. L. Fellingham, "A stochastic approach to ultrasonic tissue characterization," Ph.D. dissertation, Stanford Univ., CA., 1979.
- [20] L. L. Fellingham and F. G. Sommer, "Ultrasonic characterization of tissue structure in the *In Vivo* human liver and spleen," *IEEE Trans. Son. Ultrason.*, vol. 31, no. 4, pp. 418–428, 1984.
- [21] S. L. Marple, *Digital Spectral Analysis: With Applications*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1987.
- [22] R. Kuc, K. Haghkerdar, and M. O'Donnell, "Presence of cepstral peak in random reflected ultrasound signals," *Ultrason. Imaging*, vol. 8, no. 3, pp. 196–212, 1986.
- [23] I. M. Rosado-Mendez, L. C. Drehfal, J. A. Zagzebski, and T. J. Hall, "Analysis of coherent and diffuse scattering using a reference phantom," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 9, pp. 1306–1320, 2016.
- [24] G. Georgiou and F. S. Cohen, "Tissue characterization using the continuous wavelet transform I: decomposition method," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 48, no. 2, pp. 355–363, 2001.

- [25] K. D. Donohue, F. Forsberg, C. V. Piccoli, and B. B. Goldberg, "Analysis and classification of tissue with scatterer structure templates," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 46, no. 2, pp. 300–310, 1999.
- [26] K. Donohue, L. Huang, T. Burks, F. Forsberg, and C. Piccoli, "Tissue classification with generalized spectrum parameters," *Ultrasound in Med. & Biol.*, vol. 27, no. 11, pp. 1505 – 1514, 2001.
- [27] N. Rubert and T. Varghese, "Mean scatterer spacing estimation using multi-taper coherence," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 60, no. 6, pp. 1061–1073, 2013.
- [28] E. Feleppa, F. Lizzi, D. Coleman, and M. Yaremko, "Diagnostic spectrum analysis in ophthalmology: a physical perspective," *Ultrasound Med. Biol.*, vol. 12, no. 8, pp. 623–631, 1986.
- [29] W. Liu, J. A. Zagzebski, M. A. Kliewer, T. Varghese, and T. J. Hall, "Ultrasonic scatterer size estimation in liver tumor differentiation," *Med. Phys.*, vol. 34, p. 2597, 2007.
- [30] M. F. Insana, T. J. Hall, J. G. Wood, and Z. Yan, "Renal ultrasound using parametric imaging techniques to detect changes in microstructure and function," *Invest. Radio.*, vol. 28, no. 8, pp. 720–725, 1993.
- [31] T. J. Hall, M. F. Insana, L. A. Harrison, and G. G. Cox, "Ultrasonic measurement of glomerular diameters in normal adult humans," *Ultrasound Med. Biol.*, vol. 22, no. 8, pp. 987–997, 1996.
- [32] M. F. Insana, "Modeling acoustic backscatter from kidney microstructure using an anisotropic correlation function," *J. Acoust. Soc. Am.*, vol. 97, no. 1, pp. 649–655, 1995.
- [33] E. J. Feleppa, T. Liu, A. Kalisz, M. C. Shao, N. Fleshner, V. Reuter, and W. R. Fair, "Ultrasonic spectral-parameter imaging of the prostate," *Int. J. Imaging Syst. Technol.*, vol. 8, no. 1, pp. 11–25, 1997.

- [34] S. L. Bridal, P. Fornès, P. Bruneval, and G. Berger, “Parametric (integrated backscatter and attenuation) images constructed using backscattered radio frequency signals (25–56 MHz) from human aortae *in vitro*,” *Ultrasound Med. Biol.*, vol. 23, no. 2, pp. 215–229, 1997.
- [35] R. H. Silverman, R. Folberg, H. C. Boldt, H. O. Lloyd, M. J. Rondeau, M. G. Mehaffey, F. L. Lizzi, and D. J. Coleman, “Correlation of ultrasound parameter imaging with microcirculatory patterns in uveal melanomas,” *Ultrasound Med. Biol.*, vol. 23, no. 4, pp. 573–581, 1997.
- [36] T. Liu, F. L. Lizzi, R. H. Silverman, and G. J. Kutcher, “Ultrasonic tissue characterization using 2-D spectrum analysis and its application in ocular tumor diagnosis,” *Med. Phys.*, vol. 31, no. 5, pp. 1032–1039, 2004.
- [37] J. Mamou, A. Coron, M. Hata, J. Machi, E. Yanagihara, P. Laugier, and E. J. Feleppa, “Three-dimensional high-frequency characterization of cancerous lymph nodes,” *Ultrasound Med. Biol.*, vol. 36, no. 3, pp. 361–375, 2010.
- [38] R. Golub, R. Parsons, B. Sigel, E. Feleppa, J. Justin, H. A. Zaren, M. Rorke, J. Sokil-Melgar, and H. Kimitsuki, “Differentiation of breast tumors by ultrasonic tissue characterization,” *J. Ultrasound Med.*, vol. 12, no. 10, pp. 601–608, 1993.
- [39] R. Lavarello and M. Oelze, “Quantitative ultrasound estimates from populations of scatterers with continuous size distributions: Effects of the size estimator algorithm,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 59, no. 9, pp. 2066–2076, 2012.
- [40] A. C. Luchies, G. Ghoshal, W. D. O’Brien Jr., and M. L. Oelze, “Quantitative ultrasonic characterization of diffuse scatterers in the presence of structures that produce coherent echoes,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 59, no. 5, pp. 893–904, 2012.
- [41] A. C. Luchies and M. L. Oelze, “Backscatter coefficient estimation using tapers with gaps,” *Ultrason. imaging*, vol. 37, no. 2, pp. 117–134, 2015.

- [42] M. L. Oelze, W. D. O'Brien Jr., J. P. Blue, and J. F. Zachary, "Differentiation and characterization of rat mammary fibroadenomas and 4T1 mouse carcinomas using quantitative ultrasound imaging," *IEEE Trans. Med. Imaging*, vol. 23, no. 6, pp. 764–771, 2004.
- [43] M. F. Insana, R. F. Wagner, D. G. Brown, and T. J. Hall, "Describing small-scale structure in random media using pulse-echo ultrasound," *J. Acoust. Soc. Am.*, vol. 87, no. 1, pp. 179–192, 1990.
- [44] F. Schaefer, I. Heer, P. Schaefer, C. Mundhenke, S. Osterholz, B. Order, N. Hofheinz, J. Hedderich, M. Heller, W. Jonat, *et al.*, "Breast ultrasound elastography-results of 193 breast lesions in a prospective study with histopathologic correlation," *European journal of radiology*, vol. 77, no. 3, pp. 450–456, 2011.
- [45] K. A. Wear, R. F. Wagner, and B. S. Garra, "A comparison of autoregressive spectral estimation algorithms and order determination methods in ultrasonic tissue characterization," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 42, no. 4, pp. 709–716, 1995.
- [46] Z. Zhou, W. Wu, S. Wu, K. Jia, and P.-H. Tsui, "A review of ultrasound tissue characterization with mean scatterer spacing," *Ultrasonic imaging*, vol. 39, no. 5, pp. 263–282, 2017.
- [47] J. A. Jensen, "A model for the propagation and scattering of ultrasound in tissue," *J. Acoust. Soc. Amer.*, vol. 89, no. 1, pp. 182–190, 1991.
- [48] U. R. Abeyratne, A. P. Petropulu, and J. M. Reid, "Higher order spectra based deconvolution of ultrasound images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 42, no. 6, pp. 1064–1075, 1995.
- [49] C. Yu, C. Zhang, and L. Xie, "A blind deconvolution approach to ultrasound imaging," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 59, no. 2, pp. 271–280, 2012.

- [50] J. Ng, R. Prager, N. Kingsbury, G. Treece, and A. Gee, "Modeling ultrasound imaging as a linear, shift-variant system," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 53, no. 3, pp. 549–563, 2006.
- [51] M. Miyoshi and Y. Kaneda, "Inverse filtering of room acoustics," *IEEE Trans. Acoust., Speech Signal Process.*, vol. 36, no. 2, pp. 145–152, 1988.
- [52] J. A. Jensen, "Field: A program for simulating ultrasound systems," in *10th Nordic-Baltic Conference on Biomedical Imaging*, vol. 4, no. 1, pp. 351–353.
- [53] K. Nam, J. A. Zagzebski, and T. J. Hall, "Quantitative assessment of in vivo breast masses using ultrasound attenuation and backscatter," *Ultrason. imaging*, vol. 35, no. 2, pp. 146–161, 2013.
- [54] M. K. Hasan, M. A. Hussain, S. R. Ara, S. Y. Lee, and S. K. Alam, "Using nearest neighbors for accurate estimation of ultrasonic attenuation in the spectral domain," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 60, no. 6, pp. 1098–1114, 2013.
- [55] M. K. Hasan, M. Shifat-E-Rabbi, and S. Y. Lee, "Blind deconvolution of ultrasound images using l_1 -norm-constrained block-based damped variable step-size multichannel lms algorithm," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 8, pp. 1116–1130, 2016.
- [56] Z. Klimonda, M. Postema, A. Nowicki, and J. Litniewski, "Tissue attenuation estimation by mean frequency downshift and bandwidth limitation," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 8, pp. 1107–1115, 2016.
- [57] R. J. Gledhill, "Methods for investigating conformational change in biomolecular simulations." Ph.D. dissertation, School Chem., Southampton Univ., 2004.
- [58] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: a noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 01, pp. 1–41, 2009.

- [59] Y. Gao, G. Ge, Z. Sheng, and E. Sang, "Analysis and solution to the mode mixing phenomenon in emd," in *CISP'08*, vol. 5. IEEE, 2008, pp. 223–227.
- [60] N. I. Nizam, S. K. Alam, and M. K. Hasan, "EEMD domain ar spectral method for mean scatterer spacing estimation of breast tumors from ultrasound backscattered rf data," *IEEE Trans. Ultrason., Ferroelect., Freq. Control*, vol. 64, no. 10, pp. 1487–1500, 2017.
- [61] E. Khan, F. Al Hossain, S. Z. Uddin, S. K. Alam, and M. K. Hasan, "A robust heart rate monitoring scheme using photoplethysmographic signals corrupted by intense motion artifacts," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 3, pp. 550–562, 2016.
- [62] G. Georgiou and F. S. Cohen, "Statistical characterization of diffuse scattering in ultrasound images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 45, no. 1, pp. 57–64, 1998.
- [63] T. Varghese and J. Ophir, "Estimating tissue strain from signal decorrelation using the correlation coefficient," *Ultrasound Med. Biol.*, vol. 22, no. 9, pp. 1249–1254, 1996.
- [64] C. Simon, R. Seip, and E. S. Ebbini, "Estimation of mean scatterer spacing based on autoregressive spectral analysis of pre-filtered echo data," in *Proc. IEEE Ultrasonics Symp.*, vol. 2. IEEE, 1995, pp. 1153–1156.
- [65] R. Jirik and T. Taxt, "Two-dimensional blind bayesian deconvolution of medical ultrasound images," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 55, no. 10, 2008.
- [66] O. Michailovich and D. Adam, "Phase unwrapping for 2-d blind deconvolution of ultrasound images," *IEEE Trans. Med. Imaging*, vol. 23, no. 1, pp. 7–25, 2004.
- [67] N. E. Huang and N. O. Attoh-Okine, *The Hilbert-Huang transform in engineering*. CRC Press, 2005.

- [68] M. K. Hasan, M. S. Apu, and M. K. I. Molla, "A robust method for parameter estimation of ar systems using empirical mode decomposition," *Signal, image and video processing*, vol. 4, no. 4, pp. 451–461, 2010.
- [69] W. A. Berg, A. G. Sechtin, H. Marques, and Z. Zhang, "Cystic breast masses and the acrin 6666 experience," *Radiologic Clinics*, vol. 48, no. 5, pp. 931–987, 2010.
- [70] M. K. Hasan, M. S. E. Rabbi, and S. Y. Lee, "Blind deconvolution of ultrasound images using l_1 -norm-constrained block-based damped variable step-size multichannel LMS algorithm," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 63, no. 8, pp. 1116–1130, 2016.
- [71] H. Kim and T. Varghese, "Attenuation estimation using spectral cross-correlation," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, no. 3, 2007.
- [72] R. Kuc, "Bounds on estimating the acoustic attenuation of small tissue regions from reflected ultrasound," *Proc. IEEE*, vol. 73, no. 7, pp. 1159–1168, 1985.
- [73] M. K. Hasan, M. A. Hussain, S. R. Ara, S. Y. Lee, and S. K. Alam, "Using nearest neighbors for accurate estimation of ultrasonic attenuation in the spectral domain," *EEE Trans. Ferroelectr. Freq. Control*, vol. 60, no. 6, pp. 1098–1114, 2013.
- [74] F. L. Lizzi, M. Astor, T. Liu, C. Deng, D. J. Coleman, and R. H. Silverman, "Ultrasonic spectrum analysis for tissue assays and therapy evaluation," *Int. J. Imaging Syst. Technol.*, vol. 8, no. 1, pp. 3–10, 1997.
- [75] P. Welch, "The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms," *EEE Trans. Audio Electroacoust.*, vol. 15, no. 2, pp. 70–73, 1967.
- [76] E. Franceschini and R. Guillermin, "Experimental assessment of four ultrasound scattering models for characterizing concentrated tissue-mimicking phantoms," *J. Acoust. Soc. Am.*, vol. 132, no. 6, pp. 3735–3747, 2012.

- [77] J. Mamou and M. L. Oelze, *Quantitative ultrasound in soft tissues*. Springer, 2013.
- [78] S. R. Ara, F. Alam, M. H. Rahman, S. Akhter, R. Awwal, and M. K. Hasan, “Bimodal multiparameter-based approach for benign–malignant classification of breast tumors,” *Ultrasound Med. Biol.*, vol. 41, no. 7, pp. 2022–2038, 2015.
- [79] W.-C. Shen, R.-F. Chang, and W. K. Moon, “Computer aided classification system for breast ultrasound based on breast imaging reporting and data system (bi-rads),” *Ultrasound Med. Biol.*, vol. 33, no. 11, pp. 1688–1698, 2007.
- [80] K. Horsch, M. L. Giger, L. A. Venta, and C. J. Vyborny, “Automatic segmentation of breast lesions on ultrasound,” *Med. Phys.*, vol. 28, no. 8, pp. 1652–1659, 2001.
- [81] S. K. Alam, E. J. Feleppa, M. Rondeau, A. Kalisz, and B. S. Garra, “Ultrasonic multi-feature analysis procedure for computer-aided diagnosis of solid breast lesions,” *Ultrason. Imaging*, vol. 33, no. 1, pp. 17–38, 2011.
- [82] R. Jain, R. Kasturi, and B. G. Schunck, *Machine vision*. McGraw-Hill New York, 1995, vol. 5.
- [83] W. Yao, B. Zhao, Y. Zhao, W. Wang, and G. Qian, “Ultrasonographic texture analysis of parenchymatous organs by the four-neighborhood-pixels algorithm: clinical experiment.” *J. Ultrason. Med.*, vol. 20, no. 5, pp. 465–471, 2001.
- [84] H. Hurst, R. Black, Y. Simaika, and L.-t. Storage, “An experimental study, constable & co,” *Ltd., London*, vol. 33, 1965.
- [85] S. Golemati, T. J. Tegos, A. Sassano, K. S. Nikita, and A. N. Nicolaides, “Echogenicity of b-mode sonographic images of the carotid artery: Work in progress,” *J. Ultrason. Med.*, vol. 23, no. 5, pp. 659–669, 2004.
- [86] K. Drukker, M. L. Giger, K. Horsch, M. A. Kupinski, C. J. Vyborny, and E. B. Mendelson, “Computerized lesion detection on breast ultrasound,” *Med. Phys.*, vol. 29, no. 7, pp. 1438–1446, 2002.

- [87] Y.-H. Chou, C.-M. Tiu, G.-S. Hung, S.-C. Wu, T. Y. Chang, and H. K. Chiang, "Stepwise logistic regression analysis of tumor contour features for breast ultrasound diagnosis," *Ultrasound Med. Biol.*, vol. 27, no. 11, pp. 1493–1498, 2001.
- [88] R.-F. Chang, W.-J. Wu, W. K. Moon, and D.-R. Chen, "Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors," *Breast cancer research and treatment*, vol. 89, no. 2, p. 179, 2005.
- [89] J. C. Russ, *The image processing handbook*. CRC press, 1999.
- [90] J. Kilday, F. Palmieri, and M. D. Fox, "Classifying mammographic lesions using computerized image analysis," *IEEE Trans. Med. Imaging*, vol. 12, no. 4, pp. 664–669, 1993.
- [91] B. S. Garra, E. I. Cespedes, J. Ophir, S. R. Spratt, R. A. Zuurbier, C. M. Magnant, and M. F. Pennanen, "Elastography of breast lesions: initial clinical results." *Radiology*, vol. 202, no. 1, pp. 79–86, 1997.
- [92] W. K. Moon, C.-S. Huang, W.-C. Shen, E. Takada, R.-F. Chang, J. Joe, M. Nakajima, and M. Kobayashi, "Analysis of elastographic and b-mode features at sonoelastography for breast tumor classification," *Ultrasound Med. Biol.*, vol. 35, no. 11, pp. 1794–1802, 2009.
- [93] W. Zhu, N. Zeng, N. Wang, *et al.*, "Sensitivity, specificity, accuracy, associated confidence interval and roc analysis with practical sas implementations," *NESUG proceedings: health care and life sciences, Baltimore, Maryland*, vol. 19, pp. 1–9, 2010.
- [94] B. W. Matthews, "Comparison of the predicted and observed secondary structure of t4 phage lysozyme," *Biochimica et Biophysica Acta (BBA)-Protein Structure*, vol. 405, no. 2, pp. 442–451, 1975.
- [95] S. Zhou, J. Shi, J. Zhu, Y. Cai, and R. Wang, "Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 688–696, 2013.

- [96] A. Nahiyan and M. K. Hasan, “Hybrid algorithm for elastography to visualize both solid and fluid-filled lesions,” *Ultrasound Med. Biol.*, vol. 41, no. 4, pp. 1058–1078, 2015.
- [97] M. S. E. Rabbi and M. K. Hasan, “Speckle tracking and speckle content based composite strain imaging for solid and fluid filled lesions,” *Ultrasonics*, vol. 74, pp. 124–139, 2017.
- [98] J.-Z. Cheng, D. Ni, Y.-H. Chou, J. Qin, C.-M. Tiu, Y.-C. Chang, C.-S. Huang, D. Shen, and C.-M. Chen, “Computer-aided diagnosis with deep learning architecture: applications to breast lesions in us images and pulmonary nodules in ct scans,” *Scientific reports*, vol. 6, p. 24454, 2016.