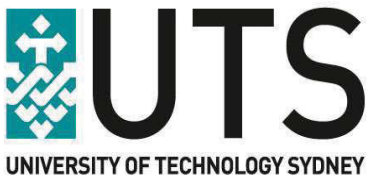

DIELECTRIC MODELLING OF HUMAN SKIN AND BREAST TISSUE IN TERAHERTZ FREQUENCIES: POTENTIAL APPLICATION TO CANCER DETECTION

CAO QUANG BAO TRUONG

THIS THESIS IS SUBMITTED
FOR THE REQUIREMENT OF DOCTOR OF PHILOSOPHY TO
FACULTY OF ENGINEERING & INFORMATION TECHNOLOGY
UNIVERSITY OF TECHNOLOGY SYDNEY

December 2015

Acknowledgments

First of all, I would like to show my gratitude to my principal supervisor, Professor Tuan Hoang. Without his great support and enthusiastic supervision, my PhD study would not have been completed.

I am also deeply grateful to Professor Hung Nguyen, my co-supervisor, for his continuous encouragement, inspiration and valuable advice.

I would like to thank my research collaborators, Associate Professor Vincent Wallace and Associate Professor Anthony Fitzgerald for providing data and professional input into my PhD research.

Special thanks is extended to send my special thanks to my friends and colleagues in the Center for Health Technologies, University of Technology, Sydney. Not only were they helpful but they made my PhD study joyful and memorable.

Last but not least, I owe my deepest gratitude to my family. They have always given me so much support and I drew strongly upon this in order to overcome the most challenging periods of my PhD candidate.

Certificate of Authorship

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of the requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature _____

Name CAO QUANG BAO TRUONG

Contents

Acknowledgments i

Certificate of Authorship ii

Contents iii

Abstract viii

List of Figures x

List of Tables xiv

Notation xvii

Author's Publications xix

1 INTRODUCTION 1

1.1 Background 2

1.2 Motivation of Thesis 4

1.3 Objectives and Contribution 7

1.4 Structure of Thesis 9

2 LITERATURE REVIEW 12

2.1 Properties of Terahertz Radiation 13

2.2 Terahertz Generation and Detection 13

2.2.1 Generation 13

2.2.2 Detection 15

2.3 Terahertz Time-Domain Spectroscopy 17

2.4 Terahertz Imaging 19

2.5 Medical Applications of Terahertz Imaging 21

2.5.1 Skin Cancer Detection 23

2.5.1.1 Overview of Skin Cancer 23

2.5.1.2 Existing Imaging Techniques of Skin Cancer Detection 24

2.5.1.3 THz Imaging for Skin Cancer 25

2.5.2 Breast Cancer Detection 29

2.5.2.1 Overview of Breast Cancer 29

2.5.2.2 Existing Techniques of Breast Cancer Detection 30

2.5.2.3 THz Imaging for Breast Cancer 33

2.5.3 Dielectric Modelling-Based Approach 35

3 PARAMETER EXTRACTION FOR THE DOUBLE DEBYE MODEL

OF HUMAN SKIN TISSUE **38**

3.1 Introduction 39

3.2 Problem Formulation 40

3.2.1 The Double Debye Model 40

3.2.2 Optimization Problem 41

3.3 Global optimization-based Algorithms 42

3.3.1 Grinding-based optimization 42

3.3.2 BB-Based optimization 44

3.3.2.1 Monotonic optimization-Based Bounding 47

3.3.2.2 Branching and Bounding Algorithm 50

3.3.3 Minimax optimization 51

3.4 Numerical Results and Discussion 53

3.5 Conclusion 62

4 THE POTENTIAL OF THE DOUBLE DEBYE PARAMETERS TO DISCRIMINATE BETWEEN BASAL CELL CARCINOMA AND NORMAL SKIN **65**

4.1 Introduction 66

4.2 Model Parameters Extraction 67

4.2.1 Examined Data 67

4.2.2 Extraction Procedure 67

4.3 Statistical Analysis 69

4.3.1 Student t-test 70

4.3.2 Normalized Percentage Difference in Mean 72

4.3.3 Correlation Analysis 74

4.3.4 ROC analysis 75

4.4 Classification Methodologies 77

4.4.1 Linear Programming 78

4.4.2 Support Vector Machine 79

4.4.3 Accuracy Estimation 80

4.5 Results and Discussion 81

4.6 Conclusion 84

5 A NEW DIELECTRIC MODEL OF HUMAN BREAST TISSUE IN TERAHERTZ REGIME 87

5.1 Introduction 88

5.2 Dielectric Properties Modelling 89

5.2.1 Dielectric Model of Human Breast Tissue 89

5.2.2 Fitting Algorithm 93

5.2.3 Data for Simulation 96

5.2.4 Fitting Results 97

5.3 Statistical Analysis 101

5.3.1 Correlation 101

5.3.2 Significance of Difference 103

5.3.3 ROC Analysis 103

5.4 Classification Potential 105

5.4.1 Support Vector Machine 105

5.4.2 Bayesian Neural Networks 105

5.4.3 Accuracy Estimation 107

5.4.4 Results 107

5.5 Discussion and Conclusion 110

6 Conclusion **113**

6.1 Summary 114

6.2 Future Research 117

Abstract

Growing developments in the generation and detection of terahertz (THz) radiation over more than two decades have created a strong incentive for researchers to study the biomedical applications of terahertz imaging. Contrasts in the THz images of various types of cancer, especially skin and breast cancer, are associated with changes in the dielectric properties of cancerous tissues. In fact, dielectric models can explain the interaction between terahertz radiation and human tissue at a molecular level just as their parameters have the potential for becoming indicators of cancer. However, dielectric modelling of various forms of human tissue remains limited due to a number of factors, especially suboptimal fitting algorithms and tissue heterogeneity.

Thanks to the high water content of human skin, its dielectric response to terahertz radiation can be described by the double Debye model. The existing fitting method using a nonlinear least square algorithm can extract the model parameters which track their measurements accurately at frequencies higher than one THz but poorly at lower frequencies. However, the majority of dielectric contrast between normal and cancerous skin tissues has been observed in the low THz range. Accordingly, this research has developed two global optimization algorithms which are capable of globally accurate tracking thereby supporting the full validity of the double Debye model in simulating the dielectric spectra of human skin in the THz frequencies. Numerical results confirm their superiority over the conventional methods. Furthermore, the next goal of the study is to apply statistical analysis to the parameters of the double Debye model in

order to test their discrimination capability of skin cancer from normal tissue. Linear programming and support vector machine algorithms have also been employed using these parameters to classify normal skin tissue and basal cell carcinoma. By combining the double Debye parameters, the classification accuracy has shown significant improvement. The encouraging outcomes confirm the classification potential of the double Debye parameters.

The double Debye model, however, has been shown to be not suitable for simulating human breast tissue due to its low water content and heterogeneous structure, thus limiting the understanding of the THz dielectric response of breast tissue. To overcome this problem, this study proposes a new non-Debye dielectric model to fit the dielectric spectra of human breast tissue. Due to the mathematical complexity of the fitting procedure, a sampling gradient algorithm of non-smooth optimization is used to optimize the fitting solution. Simulation results confirm applicability of the non-Debye model through its exceptional ability to fit the examined data. Statistical measures have also been used to analyse the possibility of using the parameters of this model to differentiate breast tumours from healthy breast tissue. Based on the statistical analysis, popular classification methods such as support vector machines and Bayesian neural network have also been applied to examine these parameters and their combinations for breast cancer classification. The obtained classification accuracies indicate the classification potential of the model parameters as well as highlighting several valuable features of the parameter combinations.

List of Figures

2.1 A THz-TDS system in transmission mode [Ferguson and Zhang, 2002]. 18

2.2 A typical time-domain imaging system using raster scan [Chan et al., 2007]. 20

2.3 A schematic of the terahertz imaging system operating in the reflection (a) and transmission (b) geometries [Pickwell-MacPherson and Wallace, 2009]. 22

2.4 Averaged THz pulses reflected off a tissue sample [Wallace, Fitzgerald, Shankar, Flanagan, Pye, Cluff and Arnone, 2004]. 27

2.5 (a) A clinical photo of an invasive nodular BCC; (b) The corresponding THz image with the maximum amplitude E_{max} for showing surface features; (c) The corresponding THz images with a normalised amplitude at $t = 2.8\ ps$ for mapping the tumour at depth [Wallace, Fitzgerald, Shankar, Flanagan, Pye, Cluff and Arnone, 2004]. 28

2.6 (a) Photomicrograph of an invasive ductal carcinoma sample; (b) the corresponding THz image using the minimum amplitude of the reflection impulses; The cancerous regions delineated from the photomicrograph (c) and the THz image (d) [Fitzgerald et al., 2006]. 33

2.7 Averaged refractive indices and absorption coefficients with the corresponding standard errors of three breast tissue types: fat, fibre, and tumour [Ashworth et al., 2009]. 34

3.1 The real and imaginary part of the relative complex permittivities of normal skin plotted in the range [0.15-1.68] THz from: the measured data of Pickwell et al. [2005], the simulated spectra using the DD models obtained by Pickwell et al. [2005] and the grinding-based algorithm . . 54

3.2 The real and imaginary part of the relative complex permittivities of BCC plotted in the range [0.15-1.68] THz from: the measurement of Pickwell et al. [2005], the simulated spectra using the DD models obtained by Pickwell et al. [2005] and the grinding-based algorithm 55

3.3 The real and imaginary part of the relative complex permittivities of normal skin: the measurement of Pickwell et al. [2005] and the simulated spectra using the DD models obtained by Pickwell et al. [2005] and the BB-GO 57

3.4 The real and imaginary part of the relative complex permittivities of BCC: the measurement of Pickwell et al. [2005] and the simulated spectra using the DD models obtained by Pickwell et al. [2005] and the BB-GO 58

3.5 The real and imaginary part of the relative complex permittivities of normal skin (case 266) of Wallace et al. [2006]: the measurement and the simulated spectra using the DD models obtained by the BB-GO . . 59

3.6 The real and imaginary part of the relative complex permittivities of BCC (case 266) of Wallace et al. [2006]: the measurement and the simulated spectra using the DD models obtained by the BB-GO 60

3.7 The real part of the relative complex permittivities of normal skin and BCC (average) from [Wallace et al., 2006]: the measurement and the simulated spectra using the DD models obtained by the BB-GO and the BB-MNO 62

3.8 The imaginary part of the relative complex permittivities of normal skin and BCC (average) from [Wallace et al., 2006]: the measurement and the simulated spectra using the DD models obtained by the BB-GO and the BB-MNO 63

3.9 The magnitude of the relative complex permittivities of normal skin and BCC in the average case of Wallace et al. [2006] plotted by: the measurement and simulations by the BB-GO and the BB-MNO 64

4.1 The p-values for comparing normal skin and BCC by using the real and imaginary part of the measured complex permittivity ϵ_m over the frequency range of 0.2-1.8 THz. The dotted horizontal red line indicates a p-value equal to 0.05. 71

4.2 The p-values for comparing normal skin and BCC by using the values of $\Im(\epsilon_{\mathrm{DD}})$, $\Im(D_1)$ and $\Im(D_2)$ respectively over the frequency range of 0.2-1.8 THz. The dotted horizontal red line indicates a p-value equal to 0.05. 72

4.3 Normalized percentage difference between normal and cancerous skin tissues in mean of the measured absorption coefficient, the refractive index, and complex permittivity. 73

4.4 The Pearson correlations of $\Im(D_1)$, $\Im(D_2)$, $\Im(\epsilon_{\mathrm{DD}})$ and $\Im(\epsilon_m)$ with the tumour percentage content in skin samples. 75

4.5 ROC curves of ε_s, ε_2, ε_∞, τ_1 and τ_2. The dotted line in red highlights the no-discrimination line. 76

4.6 The areas under the ROC curves of $\Im(\epsilon_m)$, $\Im(\epsilon_{\mathrm{DD}})$ and $\Im(D_1)$ over the frequency range $(0.2 - 1.8)$ THz. 77

5.1 The average complex permittivities with the corresponding error bars representing 95% confidence intervals for breast tumour, healthy fibrous breast tissue, and healthy fat (adipose) breast tissue from [Ashworth et al., 2009] 91

5.2 The normalized percentage difference in the average complex permittivities between the two healthy breast tissue groups (fibrous and fat) and the breast tumour. 92

5.3 The measured average complex permittivities of healthy fat tissue from [Ashworth et al., 2009] and its fitting by the THz-nDB model. 98

5.4 The measured complex permittivities of healthy fibrous breast tissue from [Ashworth et al., 2009] and its fitting by the THz-nDB model. . . 99

5.5 The measured complex permittivities of a breast tumour from [Ashworth et al., 2009] and its fitting by the THz-nDB model. 100

5.6 ROC plot for various thresholds of σ. 104

List of Tables

3.1 The double Debye parameters extracted from the measured complex permittivities of the skin samples (NS-1 and BCC-1 from [Pickwell et al., 2005]; NS(Case 266), BCC(Case 266), NS(Average), BCC(Average) from [Wallace et al., 2006]) by the grinding-based algorithm and BB-GO, and those from these references. 56

3.2 Values of R^2_{adj} corresponding to fitting the real and imaginary parts of the measured complex permittivities (NS-1 and BCC-1 from [Pickwell et al., 2005]; NS(Case 266), BCC(Case 266), NS(Average), BCC(Average) from [Wallace et al., 2006]) are compared between the BB-GO and NLS. . . 61

3.3 p-value of the Breusch-Pagan and Koenker test for the real and imaginary parts of the extracted DD models using the BB-GO for the skin samples (NS-1 and BCC-1 from [Pickwell et al., 2005]; NS(Case 266), BCC(Case 266), NS(Average), BCC(Average) from [Wallace et al., 2006]). 61

4.1 The double Debye parameters. 69

4.2 The p-values for the double Debye parameters to illustrate the significance of difference between normal skin and BCC samples. 70

4.3 The normalized percentage difference in means of the double Debye parameters between the normal skin and BCC samples. 73

4.4 The Pearson's coefficient of correlation between each of the double Debye parameters and the sample tumour content. 74

4.5 The AUCs for the five DD parameters and the AUCs averaged over the entire frequency range from 0.2 to 1.8 THz for $\Im(D_1)$, $\Im(D_2)$, $\Im(\epsilon_{\mathrm{DD}})$, $\Im(\epsilon_m)$ 77

4.6 The estimated accuracies (%) with their standard deviations by LOOCV and RRSCV for applying the double Debye parameters with the LP algorithms to classify the normal skin and BCC samples. 82

4.7 The estimated accuracies (%) with their standard deviations by LOOCV and RRSCV for applying the double Debye parameters with the SVM algorithm to classify the normal skin and BCC samples. 83

5.1 The extracted parameters of the THz-nDB model (5.5) by fitting it to averaged complex permittivities of three types of human breast tissue. . 97

5.2 The group average of the extracted parameters with their standard errors and the corresponding average R^2_{adj} of the fittings for all 74 data samples. 101

5.3 Pearson correlation between the parameters of the THz-nDB model (5.5) and the percentage of tissue components 101

5.4 P-value of independent two-tailed sample t-test for each pair of sample groups 103

5.5 AUROC values obtained by ROC of each THz-nDB parameter of (5.5) 104

5.6 SVM Classification Accuracy (%) in LOOCV corresponding to each THz-nDB parameter. 107

5.7 The estimated accuracies (%) by LOOCV and RRSCV for applying the double Debye parameters with the SVM to classify the healthy breast tissue and the breast tumour. 108

5.8 The estimated accuracies (%) by LOOCV and RRSCV for applying the double Debye parameters with the BNN to classify the healthy breast tissue and the breast tumour. 109

Notation

THz	Terahertz
NMSC	Nonmelanoma Skin Cancer
BCC	Basal Cell Carcinoma
SCC	Squamous Cell Carcinoma
MMS	Moh's micrographic surgery
BCS	Breast conserving surgery
NS	Normal skin
$A \succeq 0$	A is A positive semi-definite matrix
$A \geq 0$	A is positive, i.e all its entries are positive
$x \geq p$	$x_i \geq p_i$ with $x \in R^n$, $p \in R^n$, $i = 1, 2, ..., n$
$x \in [p, q]$	$x_i \in [p_i, q_i]$ with $x \in R^n$, $p \in R^n$, $p \in R^n$, $i = 1, 2, ..., n$
$x \in R^n_+$	$\{x \in R^n \ : \ x_i \geq 0, \ i = 1, 2, ..., n\}$
$\tilde{x}$	x is a complex variable
$\Re(x)$	Real part of a complex variable x
$\Im(x)$	Imaginary part of a complex variable x
DD	Double Debye
THz-nDB	The non-Debye model for the human breast in the terahertz regime
NLS	Nonlinear Least Square
BB-GO	Global optimisation based on the branching and bounding method
BB-MNO	Minimax optimisation based on the branching and bounding method

QP	Quadratic program
R^2_{adj}	Adjusted R^2 indicates the goodness-of-fit
GOF	Goodness-of-fit
SVM	Support vector machine
HN	Havriliak-Negami relationship
CC	Cole-Cole equation
AVR	Average value
SE	Standard error
ROC	Receiver operating characteristic
AUROC	Area under the ROC curve
TPR	True positive rate
FPR	False positive rate
CV	Cross-validation
LOOCV	Leave-one-out cross-validation
RRSCV	Repeated random-subsampling cross-validation
RBF	Radial basic function
FDTD	Finite difference time domain
BNN	Bayesian neural network
NN	Neural network

Author's Publications

The contents of this thesis are based on the following papers that have been published, accepted, or submitted to peer-reviewed journals and conferences.

Journal Papers:

1. Bao C. Q. Truong, Hoang D. Tuan, H. H. Kha, and Hung T. Nguyen, "Debye parameter extraction for characterizing interaction of terahertz radiation with human skin tissue", *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 6, pp. 1528-1537, June 2013.

2. Bao C. Q. Truong, Hoang D. Tuan, Anthony J. Fitzgerald, Vincent P. Wallace, and Hung T. Nguyen, "A dielectric model of human breast tissue in terahertz regime", *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 699-707, Feb 2015.

3. Bao C. Q. Truong, Hoang D. Tuan, Anthony J. Fitzgerald, Vincent P. Wallace, and Hung T. Nguyen, "The potential of the double Debye parameters to discriminate between basal cell carcinoma and normal skin", *IEEE Transactions on Terahertz Science and Technology*, vol. 5, no. 6, pp. 990-998, Nov 2015.

Conference Papers:

1. Bao C. Q. Truong, Hoang D. Tuan, H. H. Kha, and Hung T. Nguyen,"System identification for terahertz wave's propagation and reflection in human skin", *Fourth International Conference on Communications and Electronics*, Hue, Vietnam, pp. 364-368, Aug 2012.

2. Bao C. Q. Truong, Hoang D. Tuan, H. H. Kha, and Hung T. Nguyen,"Global optimization for human skin investigation in terahertz", *34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, USA, pp. 5474-5477, Aug 2012.

3. Bao C. Q. Truong, Hoang D. Tuan, and Hung T. Nguyen,"Near-infrared parameters extraction: a potential method to detect skin cancer", *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Osaka, Japan, pp. 33 - 36, July 2013.

4. Bao C. Q. Truong, Hoang D. Tuan, Anthony J. Fitzgerald, Vincent P. Wallace, and Hung T. Nguyen,"High correlation of double Debye model parameters in skin cancer detection", *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Chicago, USA, pp. 718-721, Aug 2014.

5. Bao C. Q. Truong, Hoang D. Tuan, Anthony J. Fitzgerald, Vincent P. Wallace, Tuan Nghia Nguyen, and Hung T. Nguyen,"Breast cancer classification using extracted parameters from a terahertz dielectric model of human breast tissue", *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Milan, Italy, pp. 2804-2807, Aug 2015.