Pattern Recognition and Tomographic Reconstruction with Terahertz Signals for Applications in Biomedical Engineering

by

Xiaoxia (Sunny) Yin

Bachelor of Engineering (Industrial Electronics),
Dalian University, 1996

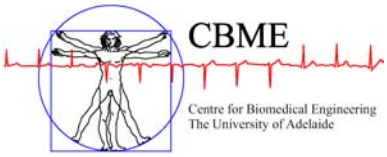
Thesis submitted for the degree of

Doctor of Philosophy

in

Electrical and Electronic Engineering
University of Adelaide

2008



Although the road ahead is endless and faraway,
I still want to pursue and explore the truth in the world

—— Xiaoxia Yin writing on January 2009 (Personal Seal)

Contents

Abstract

Over the last ten years, terahertz (THz or T-ray) biomedical imaging has become a modality of interest due to its ability to simultaneously acquire both image and spectral information. Terahertz imaging systems are being commercialized, with increasing trials performed in a biomedical setting. Advanced digital image processing algorithms are greatly need to assist screening, diagnosis, and treatment. Pattern recognition algorithms play a critical role in the accurate and automatic process of detecting abnormalities when applied to biomedical imaging. This goal requires classification of meaningful physical contrast and identification of information in images, for example, distinguishing between different biological tissues or materials. T-ray tomographic imaging and detection technology contributes especially to our ability to discriminate opaque objects with clear boundaries and makes possible significant potential applications in both *in vivo* and *ex vivo* environments.

The Thesis consists of a number of Chapters, which can be grouped in to three parts. The first part provides a review of the state-of-the-art regarding THz sources and detectors, THz imaging modes, and THz imaging analysis. Pattern recognition forms the second part of this Thesis, which is represented via combining several basic operations: wavelet transforms and wavelet based signal filtering, feature extraction and selection, along with classification schemes for THz applications. Signal filtering in this Thesis is achieved via wavelet based de-noising. The ultrafast pulses generated terahertz time-domain spectroscopy (THz-TDS), which is demonstrated to justify their decomposition in the wavelet domain as it can provide better de-noising performance. Feature extraction and selection of the terahertz measurements rely on observed changes in pulse amplitude and phase, as well as scattering characteristics of several different types of powder samples under study. Additionally, three signal processing algorithms are adopted for the evaluation of the complex insertion loss function of such samples as lactose, mandelic acid, and dl-mandelic acid: (i) standard evaluation by ratioing the sample with the background spectra, (ii) a subspace identification algorithm, and (iii) a novel wavelet packet identification procedure. These system identification algorithms enable THz measurements to be transformed to features for THz pattern recognition. Meanwhile, a novel feature extraction method involving the use of Auto Regressive (AR) and Auto Regressive Moving Average (ARMA) models on the

wavelet transforms of measured T-ray pulse responses of *ex vivo* osteosarcoma cells as well as other biomedical materials is presented. Classification schemes are carried out via simple and robust schemes, such as the linear Mahalanobis distance classifier, and the non-linear Support Vector Machine (SVM) classifier. In particular, SVMs are used as a learning scheme to achieve the identification of two classes of RNA samples and multiple classes of powered materials. Coherent terahertz detection hardware—THz time-domain spectroscopy (THz-TDS)—is used to obtain all the data for validation of these classification schemes.

The past decade has witnessed the tremendous development of terahertz instruments for detecting, storing, analysing, and displaying images. Terahertz time-domain spectroscopy (THz-TDS) is a broadband technique that generates and detects THz radiation in a synchronous and coherent manner. By contrast, the newly developed THz quantum cascade laser is a narrow-band radiation source that provides potential for realising compact systems; they produce image data with higher average power levels. The third part of this Thesis discusses methods to improve the capability of both broad- and narrow-band terahertz imaging, driven by computer-aided analytical techniques. A wavelet based reconstruction algorithm for terahertz computed tomography is represented to show how this algorithm can be used to rapidly reconstruct the region of interest (ROI) with a reduction in the measurements of terahertz responses, compared with a standard filtered back-projection technique. These reconstruction algorithms are applied to the analysis of acquired experimental data and to locally recover the two-dimensional (2D) and three-dimensional (3D) structures of several optically opaque objects. Moreover, a segmentation technique based on two dimensional wavelet transforms is investigated for the identification of different materials from the reconstructed CT image.

Statement of Originality

This work contains no material that has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of the thesis, when deposited in the University Library, being available for loan, photocopying, and dissemination through the library digital thesis collection, subject to the provisions of the Copyright Act 1968. Copying or publication or use of this thesis or parts thereof for financial gain is not allowed without the author's written permission. Due recognition shall be given the author, and the University of Adelaide, in any scholarly use that may be made of any material in the thesis.

31st December 2008

Signed

Date

Acknowledgments

A great number of people have collaborated to make this PhD an exceptionally rewarding and memorable experience. I extend my sincerest thanks to my family, friends, colleagues, and supervisors for their support and encouragement.

I thank my supervisor Prof Derek Abbott for introducing me to the world of terahertz imaging and for the continuous flow of ideas that he provided. I also thank him for his encouragement and tireless work in reviewing journal publications and this Thesis. My co-supervisor Dr Brian W.-H. Ng provided welcome advice and contributed significantly to my PhD project. His work in proofing journal publications and this Thesis is also appreciated. Dr Samuel P. Mickan assisted in supervision in the early stages of this work and I would also like to thank him for his work.

I wish to thank Dr Bradley S. Ferguson (Tenix - Electronic Systems Division, now at Raytheon) for providing pulse CT data and Dr Bernd M. Fischer for providing the RNA data. Many thanks to Prof Sillas Hadjiloucas (Cybernetics, School of Systems Engineering, The University of Reading) for hosting my two month visit at Reading University and his important instruction in THz spectroscopy during this period. My thanks to Prof Lynn Gladden, (Head of Department of Chemical Engineering, University of Cambridge), and J. Axel Zeitler (Department of Chemistry, University of Cambridge) for their collaboration on the THz quantum cascade laser imaging work. In addition, I would like to express my appreciation to Prof Xi-Cheng Zhang, (J. Erik Jonsson Distinguished Professor, Director, Center for Terahertz Research at the Rensselaer Polytechnic Institute) for hosting my four month visit at Rensselaer Polytechnic Institute. His important support exposed me to an international advanced terahertz laboratory and a series of signal processing experiments.

Despite being so far away, my family in China and my friends in Adelaide have been a constant source of support. My very special appreciation goes to my husband, Cheng Shao, for being such a wonderful companion who brings so much happiness and love into my life.

Finally I am deeply grateful to the School of Electrical & Electronic Engineering at the University of Adelaide for the divisional scholarship that has allowed me to have this fantastic opportunity to live and work in Adelaide.

For personal financial support I would like to thank the Research Abroad Scholarship from The University of Adelaide (2006), the Roger Pysden Memorial Fellowship from the Australia Business, ABL State Chamber (2006), a Mutual Community Postgraduate Travel Grant from the University of Adelaide (2007), an Overseas Travel Fellowship from the Australian Research Council Nanotechnology Network (ARCNN) (2007), and a D. R. Stranks Travelling Fellowship from The University of Adelaide (2008).

Xiaoxia Yin, December 2008.

Conventions

The following conventions have been adopted in this Thesis:

1. **Definitions.** The T-ray band is defined in this Thesis to span from 0.1 to 10 THz (THz = 10^{12} Hz). This is an emerging definition in the literature. The T-ray band overlaps a little with the millimetre wave band (at lower T-ray frequencies) and the far-infrared band (at higher T-ray frequencies).

2. **Notation.** The acronyms and symbols used in this Thesis are defined in the Glossary on p. 403. Standard abbreviations are used for the order of magnitude: T, tera-, 10^{12}; G, giga-, 10^{9}; M, mega-, 10^{6}; k, kilo-, 10^{3}; m, milli-, 10^{-3}; μ, micro-, 10^{-6}; n, nano-, 10^{-9}; p, pico-, 10^{-12}; f, femto-, 10^{-15}.

3. **Units.** All T-ray electric field amplitude data in this Thesis is expressed in arbitrary units (AU). This means that the amplitude of the T-ray electric field has been normalised to a peak measurement, or to another reference measurement on the same system, depending on the experiment.

4. **Spelling.** Australian English spelling conventions have been used, as defined in the Macquarie English Dictionary (A. Delbridge (Ed.), Macquarie Library, North Ryde, NSW, Australia, 2001).

5. **Typesetting.** This document was compiled using LATEX2e.

6. **Mathematics.** MATLAB code was written using MATLAB Version 7.0. Manufacturer: the MathWorks Inc., Natick, MA, USA; URL: http://www. mathworks.com.

7. **Referencing.** The Harvard style has been adopted for referencing.

8. **Diagram colours.** The following colour scheme is approximately followed in the diagrams:

- Blue: T-ray beams (T-rays are invisible to the naked eye).
- Red: laser beams (the majority of laser beam sources appearing in the diagrams are from optical femtosecond lasers).
- Yellow: motion stages and samples (motion stages represent moving equipment, and samples can be composed of a wide variety of substances).

- Orange: electrooptic and semiconductor crystals (these are typically T-ray sources and detectors).
- Black/white: equipment, including optical components such as lenses and beam splitters.

9. **URLs.** Universal Resource Locators are provided in the Thesis for finding information on the world wide web using the hypertext transfer protocol (HTTP). The information at the locations listed was current on 31st December 2008.

Publications

Book Chapters

YIN-X. X, NG-B. W.-H, FERGUSON-B AND ABBOTT-D (2007). Wavelet based local coherent tomography with an application in terahertz imaging, in W. G. Kropatsch, M. Kampel, A. Hanbury (Eds.), *Computer Analysis of Images and Patterns, Lecture Notes in Computer Science*, Publ: Springer-Verlag, **4673**, pp. 878–885.

Journal Articles

YIN-X. X, NG-B. W.-H, FERGUSON-B, AND ABBOTT-D (2008). Wavelet based local tomographic image reconstruction using terahertz techniques, *Digital Signal Processing*, **DOI:10.1016/j.dsp.2008.06.009**.

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2007). 2-D wavelet segmentation in 3-D T-ray tomography, *IEEE Sensors Journal*, **7**(3), pp. 342–343.

YIN-X. X, NG-B. W.-H, FISCHER-B.M, FERGUSON-B AND ABBOTT-D (2007). SVM applications in terahertz feature sets for identification of RNA data, *IEEE Sensors Journal*, **7**(11-12), pp. 1597–1608.

YIN-X. X, KONG-K. M., LIM-J. W., NG-B. W.-H, FISCHER-B. M, FERGUSON-B AND ABBOTT-D (2007). Enhanced T-ray signal classification using wavelet preprocessing, *Medical & Biological Engineering & Computing (*Springer), **45**(6), pp. 611–616.

YIN-X. X, NG-B. W.-H, FERGUSON-B, HADJILOUCAS-S AND ABBOTT-D (2007). Application of auto-regressive models and wavelet sub-bands for classifying terahertz pulse measurements, *Journal of Biological Systems*, **15**(4), pp. 551–571.

WITHAYACHUMNANKUL-W, PNG-G, **YIN-X. X**, JONES-I, LIN-H, UNG-B, BALAKRISHNAN-J, NG-B. W.-H., FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2007). T-ray sensing

and imaging, *Proceedings IEEE*, **95**(8), pp. 1528–1558.

Conference Articles

YIN-X. X, NG-B. W.-H., ZEITLER-J. A., NGUYEN-K. L., GLADDEN-L., AND ABBOTT-D. (2008). Local reconstruction for three dimensional therahertz imaging using a CW quantum cascade laser, *Proceedings of the 2008 International Conference on Image Processing, Computer Vision, & Pattern Recognition*, 14-17 July, 2008, Las Vegas, Nevada, USA, pp. 252–258. [Invited oral paper].

FISCHER-B. M., **YIN-X. X**, NG-B. W.-H., ABBOTT-D., GALVÕ-R. K. H., PAIVA-H. M., HADJILOUCAS-S., WALKER-G. C., AND BOWEN-J. W. (2008). Subspace and wavelet-packet algorithms for de-noising and classifying broadband THz transients, *33rd International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)*, 15-19 September, 2008, California Institute of Technology, Pasadena, California, USA, **DOI:10.1109/ICIMW.2008.4665758**.

YIN-X. X, HADJILOUCAS-S, FISCHER-B. M, PAIVA-H. M, GALVÕ-R. K. H, NG-B. W.-H, WALKER-G. C, BOWEN-J. W AND ABBOTT-D (2007). Classification of lactose and mandelic acid THz spectra using subspace and wavelet-packet algorithms, *Proceeding of SPIE Microelectronics, MEMS, and Nanotechnology*, **6798**, Article Number 679814.

YIN-X. X, NG-B. W.-H, FERGUSON-B AND ABBOTT-D (2007). Terahertz local tomography using wavelets, *IEEE Joint 32nd International Conference on Infrared and Millimeter Waves and 15th International Conference on Terahertz Electronics (IRMMW-THz 2007)*, 2-5 September, 2007, Cardiff, UK, pp. 514–515.

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2007). 2D wavelet-based segmentation by fusion in 3D terahertz tomography, *IEEE International Symposium on Industrial Electronics (ISIE)*, 4-7 June, 2007, Vigo, Spain, pp. 3409-3414. [This paper won an IEEE (ISIE) student best paper].

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2006). Feature extraction from terahertz pulses for classification of RNA data via support vector machines, *Proceeding of SPIE Micro- and Nanotechnology: Materials, Processes,*

Packaging, and Systems III, 11-13 December 2006, Adelaide, Australia, **6415**, Article Number 641516.

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S.P AND ABBOTT-D (2006). Wavelet based segment detection and feature extraction for 3D T-ray CT pattern classification, *IEEE Signal Processing Society: 12th Digital Signal Processing Workshop 4th signal Processing Education Workshop: Filtering*, September 24-27, 2006, Wyoming, USA, pp. 602–607.

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2006). Information fusion and wavelet based segment detection with applications to the identification of 3D target T-ray CT imaging, *IEEE Joint 31st International Conference on Infrared and Millimeter Waves and 14th International Conference on Terahertz Electronics (IRMMW-THz 2006): Biological and Medical Applications*, September 18-22, 2006, Shanghai, China, p. 187.

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2006). Statistical model for the classification of the wavelet transforms of T-ray pulses, *IEEE International Conference on Pattern Recognition (ICPR 2006): Advances in Basic Methodology II*, 20-24 August 2006, Hong Kong, China, **III**, pp. 236–239. [This paper won The International Association for Pattern Recognition (IAPR) Hong Kong travel stipend].

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2006). Classification of human osteosarcoma cells via wavelet transform, and AR parametric modeling of T-ray pulsed signals, *Wavelets and Applications Conference (WavE 2006)*, July 10-14 2006, Lausanne, Switzerland, p. A-80. [This paper won a WavE Swiss travel fellowship].

YIN-X. X, NG-B. W.-H, FERGUSON-B, MICKAN-S. P AND ABBOTT-D (2005). One dimensional wavelet transforms and their application to T-ray pulsed signal identification, *Proceeding of SPIE Photonics: Design, Technology and Packaging II*, December 11-15, 2005, Brisbane, Australia, **6038**, Article Number 603829.

List of Figures

List of Tables