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

Chapter 1

Introduction and Motivation

TERAHERTZ pattern recognition is a signal processing procedure for extracting information from THz spectroscopic data. This is an area of study that is intimately tied to the measurement of terahertz data sets.

A challenge for THz pattern recognition systems is to explore the available methods for the identification of specific materials based on the THz responses. Most molecules show very complicated THz absorption spectra with a multitude of absorption lines. Those absorption lines are subject to thermal broadening at room temperature and are produced by classical electromagnetic interactions (Born and Wolf 2002). Some of the traditional spectral analysis methods, e.g. infrared spectroscopy, have greater Rayleigh scattering and are unable to access lower frequency vibrational modes, which motivates the current THz work.

The Thesis describes several novel pattern recognition algorithms, which are applied to measured THz data. An algebraic reconstruction algorithm based on wavelet transforms of a local image is also presented. Meanwhile, since THz sources and detectors have developed to a point where high signal-to-noise ratio and reasonable acquisition rates are possible, the combination of two dimensional and three dimensional imaging with spectroscopic information capture is an important elaboration of the current THz imaging system.

1.1 Introduction

This chapter introduces the field of THz (T-ray) spectroscopy and discusses the motivation for this work towards THz imaging and identification systems. It provides a road map for the Thesis and a concise summary of the novel contributions represented by this work.

1.2 Background

1.2.1 Terahertz radiation

The terahertz (THz) part of the electromagnetic spectrum lying between the millimetre wave and infrared (100 GHz-10 THz) is of significant importance to the biological sciences, because complementary information to traditional spectroscopic measurements on low-frequency bond vibrations, hydrogen bond stretching and torsions in liquids and gases may be obtained. The vibrational spectral characteristics of bio-molecules, which lie in this range (wavenumbers between 3.3-333 cm^{-1}) make T-rays a promising sensing modality for clinical diagnosis. Since THz photons, (or T-rays), have significantly lower energies (e.g., only 0.04 MeV at 100 GHz) than X-rays, they have been considered by many as non-invasive. Although non-linear interactions between biological tissue and coherent THz radiation have been predicted by Fröhlich (1980) and experimentally verified by the careful work of Grundler and the analysis of Kaiser (Grundler and Kaiser 1992) in the '90s, the widely held view at the moment is that any measurement technique that operates at THz frequencies should be considered as non-invasive. Such a conclusion is based on the assumption that in the absorption processes involved when THz pulses interact with biological tissue, the Gibbs free energy conveyed in the THz light beam is insufficient to drive chemical reactions. For example, the molar energy at a frequency f of 100 GHz would be given from $E = Nhf$ where $N = 6.023 \times 10^{23}$ mol^{-1}, (Avogadro's number), and $h = 6.626 \times 10^{-34}$ Js (Planck's constant), the calculated value of only $E = 0.04$ kJ mol^{-1} is so low (approximately 100 times lower than the amount of molar energy required for ATP hydrolysis) that for most practical purposes we may assume that the interference with biochemical processes would be minimal. Fig. 1.1 illustrates the different parameters used to describe where the T-ray frequencies, from 100 GHz to 10 THz, lie on the electromagnetic spectrum.

NOTE:
This figure is included on page 3 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 1.1. Electromagnetic spectrum. Illustration of Electromagnetic spectrum (Figure courtesy of Sillas Hadjiloucas, Department of Cybernetics, The University of Reading, Reading, UK).

Another advantage of performing imaging based on the optical properties of biological tissue with THz radiation is lower scattering than infrared light. Organ differentiation on the basis of tissue water content using microwave transmission or reflection measurements is impractical because the diffraction limited minimum spot size for a free-space beam is too large to avoid beam spill-over around most tissues and organs. From a technological point of view, THz imaging needs to compete with positron emission tomography (PET) imaging that has picomolar sensitivity but poor spatial resolution and magnetic resonance imaging (MRI) that offers millimolar sensitivity with high spatial resolution. Indeed, a diffraction limited imaging system operating at 1 THz would have a spatial resolution of 300 μm, which should be considered sufficient for many biomedical applications. Meanwhile, since 70% of the human body is composed of water, a large part of the energy in the excitation pulse is attenuated, and the resultant patterns of many biomedical samples can be identified via applying advanced signal processing techniques.

1.2.2 THz pulsed imaging and continuous wave imaging

In this Thesis, THz pulsed imaging and continuous wave (CW) THz quantum cascade laser (QCL) imaging are utilised for THz experiment data measurements. Following Karpowicz *et al.* (2005a), the comparison between the two different types of imaging is listed here to highlight their respective strengths and weaknesses.

First, we highlight the differences in hardware. Regarding the pulsed THz imaging system, a standard pump-probe set-up using a femtosecond pulse laser is used. A photoconductive antenna or a non-linear crystal through optical rectification is used as a terahertz emitter. The THz response is detected via photoconductive antenna device or electrooptic sampling driven by optical probe pulses. A continuous wave (CW) THz imaging system may apply purely electronic sources, e.g. the Gunn diode, without involving a pump-probe setup and a time delay scan. Photomixing techniques may be used in the emission and detection of terahertz radiation. Quantum cascade lasers and backward wave oscillators also allow for generation of THz radiation. In order to recover depth resolution of target measurement, a THz CW imaging system does this at the cost of decreased signal-to-noise ratio (Karpowicz *et al.* 2005a). Interferometry is used to recover phase information; and multiple or tunable sources with sufficient frequency range allow recovery of frequency-domain information.

Differences in data acquisition between CW and pulsed systems motivate us to develop different algorithms for tomographic reconstruction. Pulsed THz time domain imaging affords data richer in information but may require more advanced processing techniques. Depth information can be recovered from pulse timing. The temporal location of the reflected peak indicates a change in the optical path length, and the time delay between these reflected pulses and the main pulse can easily be converted to depth. For a CW THz imaging system, in the form of a raster, intensity image data are yielded if only a fixed-frequency source and a single detector are involved.

1.3 Outline of Thesis

Chapter 1 introduces T-ray biosensing and the key contributions presented in this Thesis, while Chapter 2 reviews the literature on T-ray generation and detection. Chapter 3 reviews terahertz imaging modes, which are the platforms to achieve THz pattern recognition, which is presented in Chapter 4. These Chapters provide a survey of the

state of most current research in the THz field. These reviews offer new perspectives on THz issues, and identify areas in need of further research, which form the motivation of this Thesis.

Chapter 5 develops a pattern recognition framework for the identification of materials in pulsed THz images. Chapter 6 presents a review of the theory and practical techniques regarding the application of wavelets to signal and image analysis. Wavelet soft threshold denoising is also demonstrated in this Chapter, which attempts to isolate material information present in the THz waveforms from systematic and random noise data. It is shown that this contributes to improvements in THz pattern recognition system performance. Chapter 7 describes five feature extraction algorithms for the resolution of different THz pattern recognition applications. In particular, ARMA models for wavelet subband coefficients form one of a few novel contributions in this Thesis. The final step for pattern recognition introduces the classifier, which is discussed in Chapter 8. Several popular classification schemes are reviewed in this Chapter. A number of THz case studies concerning THz pattern recognition with application of the represented pattern recognition framework and relative algorithms are illustrated in Chapter 9.

Radon transforms for tomographic reconstruction are introduced in Chapter 10. This Chapter combines traditional computed tomography (CT) with conventional THz imaging. The reconstruction formulae are expressed in both the time and frequency domains. A novel segmentation technique using 2D wavelet transforms is illustrated in Chapter 11, where polyethylene samples are used to achieve reconstructed tomographic images and extracted segments corresponding to different media. Chapter 12 highlights a local tomographic reconstruction of THz measurements using wavelets, with aims to reduce measurement time in practical application. Chapter 13 further applies the local reconstruction algorithm on CW THz QCL imaging. The optical properties of the QCL on a cubic sample with complex contours are explored.

In the case of pulsed THz tomography our rationale for the use of wavelets is that both THz pulses and wavelets are time localised—this matching results in a small number wavelet coefficients with high signal power, whereas noise is non-localised and thus noise power is thinly spread over many wavelet coefficients. However, in the present chapter where we adopt CW THz tomography, the rationale is mainly that short wavelet filters have good 2D localisation in the computed DWT, which should

limit the spatial extent of reconstruction artifacts. As in the CW case, we are considering *local* reconstruction, we aim to demonstrate that the wavelet localization feature is advantageous for zooming into the finer details for local reconstruction.

Chapter 14 concludes the Thesis with a summary of its outcomes and conclusions, and recommendations for extending T-ray pattern recognition and computed tomographic reconstruction in future work.

Appendix A introduces the oblique projection operation. It is important to understand subspace system identification algorithms. Appendix B provides further details about back projection algorithms. Appendix C details on error analysis with respect to wavelet based local reconstruction. Appendix D contains the details and specifications of the components of both the pulsed THz imaging hardware and continuous wavelet terahertz imaging QCL hardware, utilised in this Thesis. Appendix E lists some of the key algorithms for the realisation of THz pattern recognition tasks and computed tomographic reconstruction. These are used for THz data analysis and the code is written using Matlab software. These algorithms are also found in the attached CD-ROM, entitled X. X. Yin PhD Thesis/Matlab Algorithms.

The basic structural interconnections of the main Chapters are shown in Fig. 1.2. This flow chart shows which Chapters contain original experimental work, and how they are linked. The Thesis progresses along this chart from top to bottom and left to right.

1.4 Original contributions

This Thesis makes a number of significant contributions to the body of THz science and technology.

On the material identification front, a classification framework for THz spectroscopy is proposed. This framework encompasses preprocessing, feature extraction and classification techniques. The denoising of terahertz signals by modifying wavelet transform coefficients (discrete wavelet transforms and discrete wavelet packet transforms) is investigated experimentally (Section 6.5 and Section 9.2.4). Three feature extraction algorithms are developed. The first uses statistical modeling (AR/ARMA) using wavelet coefficients to extract highly descriptive features, represented in Section 7.4. The second uses oriented frequency components as classification features, illustrated in Subsection 9.5.2. The third is to apply system identification for feature extraction,

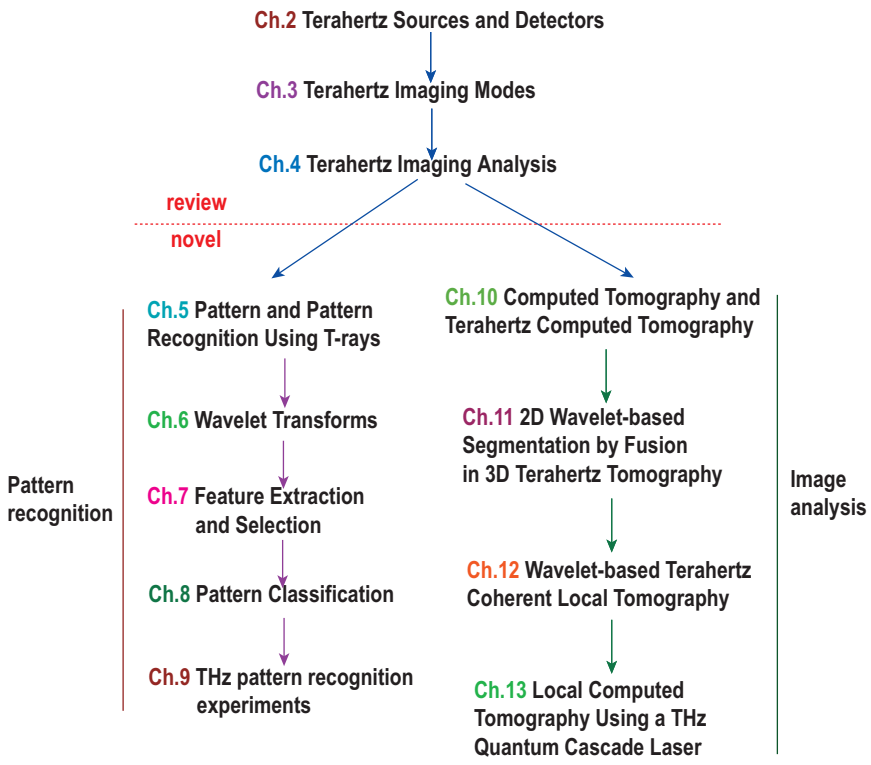


Figure 1.2. Thesis structural flow chart. This flow chart indicates how the Chapters in the Thesis fit together. Starting with the literature review, Chapters above the dotted line survey existing work. The original contributions are found in the Chapters below the dotted line. They are separated into two parts. One relates to the current pattern identification system, and the other is associated with the image analysis research represented in this Thesis. The review Chapters summarise the state-of-the-art and provide the foundation for the experimental Chapters.

consisting of subspace and wavelet packet algorithms (Section 7.5), which is the result of research in collaboration with The University of Reading. Three classifiers are applied for the final evaluation of classification, including Mahalanobis classifier (Section 8.1), the Euclidean discrimination matric (Section 8.2), and the support vector machine (Section 8.3).

Finally, several experimental case studies are conducted to verify the performance of the classification framework. The identification of cellular differences between normal human bone cells and human osteosarcoma cells is demonstrated in Section 9.4.1, 2-class RNA patterns & multiclass terahertz pulses are studied in Section 9.5, and the identification of lactose and mandelic acid THz spectra, which are overviewed in Section 9.3.

The goal of THz material identification via a classification framework, is to achieve a quantitative representation of THz spectroscopy. The goal of THz computed tomographic reconstruction is to achieve a description of the complete target of interest (Duda *et al.* 2001) via applying T-ray imaging measurements of the whole object. Chapter 11 represents a novel segmentation technique, which is derived based on THz pulsed (i.e. broadband) computed tomographic reconstruction. According to the fact that, with the incremental wavelet scale, noise is reduced and the target intensity (energy) is increased in an image, it turns out that, after extracting the low frequency 2D subband of fused T-ray CT images, an increased signed energy with an increase in wavelet scale is used as a cue to extract the target regions.

Terahertz Computed Tomography (THz-CT) is a form of coherent tomography. For time-domain terahertz measurements, it is impractical to achieve full exposure data for only local reconstruction of the subregion of interest. In order to overcome the ill-posed inverse problem from traditional filtered back projection algorithms, we apply a wavelet-based algorithm to reconstruct THz-CT images with a significant reduction in the required measurements when the region of interest is small. The algorithm recovers an approximation of the region of interest from terahertz measurements within its vicinity, and thus improves the feasibility of using terahertz imaging to detect defects in solid materials and diagnose disease states for clinical practice, etc. This is presented in Chapter 12.

In collaboration with The University of Cambridge, the local reconstruction of the region-of-interest (ROI) from a 3D terahertz imaging is obtained via a quantum cascade laser (QCL), illustrated in Chapter 13. It is an important step in the understanding wavelet based techniques and traditional filtered back projection (FBP) to map terahertz local measurements for resolution. In this experiment, a clown's head (made of polystyrene), with a hole inside with complex contours, is used as a sample to explore the ability of a QCL to image target object with complex contours. The resultant analysis has shown that wavelet-based reconstructions offer robust reconstruction performance in the local reconstructed shape of the target. Fuzzy c-means (FCM) segmentation (Chuang *et al.* 2006) is used in post-processing to make measurements on the target's internal structure, which provides an iterative measurement to minimize the classification error allowing each pixel to be the member of all the possible classes with varying membership.

Perhaps the most significant contribution of this Thesis is to realise wavelet based T-ray local computed tomography. From the optical design point of view, this technique is validated via both THz broadband (pulsed) and narrow band (CW) tomographic imaging schemes. The latter is the result in collaboration with The University of Cambridge, with the use of an advanced THz quantum cascade laser emitting at 2.9 THz. The THz QCL is a GaAs-AlGaAs bound-to-continuum superlattice design and mounted on the cold finger of a continuous-flow helium-cooled cryostat. By contrast with THz pulsed imaging, QCLs afford high power sources of continuous wave (CW) radiation to achieve deeper penetration of samples. From the mathematical derivation point of view, the development of reconstruction algorithms is to combine together the separate wavelet transforms and ramp filters, which forms an improved version of traditional filtered back projection. This is because the traditional filter back projection algorithm introduces a discontinuity in the derivative of the Fourier transform at zero frequency. This means that the traditional Radon transform inversion does not allow a complement of local reconstruction of a CT image. Our wavelet based algorithm used in this Thesis realises actual *local* THz tomographic reconstruction. It turns out to be a rapid tomographic reconstruction with reduced measurement time. Meanwhile, for local tomographic reconstruction using THz pulsed imaging (broadband THz imaging setup), we consider to adopt different constants based on the different projection angles to reduce the loss of necessary information in the ROI. We obtain local tomographic reconstruction with reduced artificial noise and a clear pattern boundary.

These contributions serve to advance the goal of the development of practical 3D THz inspection systems. They provide substantial improvements over the existing state-of-the-art and serve to extend THz applications to new and potentially ground-breaking realms.

Chapter 2

Terahertz Sources and Detectors

THIS Chapter describes techniques and systems regarding terahertz sources and detectors, mainly reviewing basic concepts and principles to help understand THz imaging instruments. The terahertz technologies outlined in this Chapter provide background knowledge to the experimental work on the identification and reconstruction of terahertz spectroscopy data later in the Thesis. This Chapter also covers the imaging modes, which involve both ultrafast pulsed systems and terahertz quantum cascade lasers. The literature review for terahertz imaging analysis and imaging modes are reviewed in the Chapters 3 and 4, respectively.

2.1 Introduction

Terahertz science has become increasingly popular in the recent decade due largely to the advent of time domain spectropy (TDS) with ultrashort-pulse laser sources, which makes possible for researchers to carry out time-resolved 'far-infrared' (FIR) studies and to explore spectroscopy and imaging applications in the submillimeter wavelength regime.

Terahertz phenomena follow fundamental scientific interests. The goal of this Chapter is to review the basic principles of different types of terahertz sources and detectors that are now possible with terahertz spectroscopy. They are fundamental to the understanding and conducting of the terahertz experiments, described later in the Thesis.

This Chapter will be organized as follows. Section 2.1 briefly introduces the history of terahertz radiation. Section 2.2 describes basic solid-state lasers, which are the principle components to achieve pulsed and continuous modes of terahertz spectroscopy. The topics range from conversion of ultrafast free space laser pulses, to free-electron lasers, as well as quantum cascade lasers. Section 2.3 introduces hardware devices for generating terahertz radiation. These terahertz emitters include on electrooptic (EO) emitters; photoconductive (PC) antennas; planar antennas and magnetic field enhancement devices, along with terahertz photomixing. Section 2.4 discusses different types terahertz receivers based on different optical sampling methods. Electrooptic versus photoconductive and synchronous versus asynchronous sampling techniques are investigated.

2.2 The history of T-rays

The term *terahertz* gained popularity among spectroscopists during the mid 1970's, (Kerecman 1973, Ashley and Palka 1973, Fleming 1974, Siegel 2002). A decade later, as a result of the efforts of many researchers and scientists, advanced techniques in optical rectification and photoconduction had been developed (Auston 1983, Smith *et al.* 1988) (see Section 2.4). These techniques made it possible to produce THz radiation directly using multimode lasers, (van Exter *et al.* 1989, Cantor *et al.* 1981), such as Ti:sapphire based lasers and free-electron lasers (Section 2.3.1). In 1989, Martin van Exter introduced a terahertz time domain spectroscopy system (TDS) (van Exter *et al.* 1989). This

added momentum to the field of T-ray generation and detection and some simple research began, including the extraction of material parameters from THz-TDS measurements.

Following this, Hu and Nuss in 1995 used a traditional scanned imaging system to acquire two dimensional (2D) images. This was the beginning of geometric image formation of an object in the T-ray frequency range. Terahertz pulsed imaging (TPI) was viewed as a novel and promising method, especially applied to medical diagnostics (Löffler *et al.* 2001).

Subsequently, much effort was devoted to system improvement (Löffler *et al.* 2002). This included progress towards wavelet de-noising of terahertz pulse imaging data (Ferguson and Abbott 2001b), de-noising techniques for terahertz responses of biological samples (Duvillaret *et al.* 1996, Ferguson and Abbott 2001a), reducing measurement time (Zhao *et al.* 2002) and improving spatial (Chen *et al.* 2000) and depth resolution (Johnson *et al.* 2001). Improved systems made possible the development of new systematic approaches, such as time resolution of TPI in translation and reflection mode (Mittleman *et al.* 1997), dark-field TPI (Löffler *et al.* 2001), THz pulsed near-field image with good spatial resolution (Federici *et al.* 2002, Schade *et al.* 2004). In turn, T-ray computed tomographic (CT) technology was applied (Wu *et al.* 1996, Jiang and Zhang 1998a, Ferguson *et al.* 2002c, Ferguson *et al.* 2002b). A review of further imaging modalities can be found in Chapter 3.

In the last five years, the rapid improvement of T-ray detectors and sources resulted in many technical advantages in THz-TDS and therefore opened up new fields of application. In particular, T-ray pulsed technology was used to image opaque objects. One promising application is the inspection of biomedical tissue. Examples are the separation of tumour cells from normal tissue (Woodward *et al.* 2002) *ex vivo*, and the study of *in vivo* normal to pathological samples of human skin (Gladkova *et al.* 2000), three dimensional detection of tooth decay (Arnone *et al.* 2000), together with optical imaging and classification of a bird head (Löffler *et al.* 2002). More spectroscopy and image analysis is reviewed in Chapter 4.

Time-domain TPI has the advantage of providing a broad frequency spectrum, but it incurs the high cost of an expensive femtosecond laser. Continuous wave (CW) imaging is attractive since it is a tunable, compact, and cost effective system.

A coherent all-optoelectronic THz measurement system via photomixing techniques is normally implemented to produce tunable CW THz radiation. It was first demonstrated in 1998 by Verghese *et al.* (1998). Following this, Gu *et al.* (1999) pointed out that a tunable dual-wavelength external cavity laser diode was a promising laser source for the generation of tunable CW-THz radiation. Phase sensitive continuous-wave THz imaging using diode lasers was introduced in 2004, with image capture rates comparable with those from state-of-the-art pulsed THz systems (Gregory *et al.* 2004). These diode lasers are cost effective for producing CW THz radiation. Additionally, we also noted that in literature, Siebert *et al.* (2002) used two colour Ti:sapphire lasers to generate CW T-rays with an extension to biomedical applications. However, there exists an obvious drawback in that two expensive Ti:sapphire laser sources are required. The CW THz radiation mentioned above is generated by applying a pair of laser sources and photomixing techniques to achieve a tunable frequency range. For comparison, an entirely electronic generation and detection scheme is illustrated by Karpowicz *et al.* (2005b), and is a relatively compact CW THz imaging system with a few components. Such a system could easily be made portable, thus opening a much wider range of possible applications. However, as it lacks depth and phase information, the system only yields two-dimensional intensity images.

Photomixing techniques (in Section 2.4.3) applied to the generation of CW terahertz radiation, very severely limit the output power to $\sim 1\ \mu$W at frequencies above 1 THz (Kim *et al.* 2006), mainly because of intrinsic parasitic impedance of the device and the impedance mismatch between the device and the radiating antenna (Darmo *et al.* 2004). Quantum cascade lasers (QCLs) operating in the terahertz range, however, have remained elusive for a long time until 2004, when the first terahertz QCLs were reported (Köhler *et al.* 2001). This is because of an insuperable barrier that exists in the phonon *reststrahlen* band, which causes difficulty in the further expansion to lower emission terahertz frequencies. Further, the development of a suitable waveguide is necessary (Sirtori *et al.* 1998a). It aims to confine light of the long wavelength within terahertz range to an epilayer compatible with molecular beam epitaxy technology and with low absorption losses onto the laser mode (Tredicucci *et al.* 2005). Recent novel design concepts have led to continuous-wave operation and high output powers (Rochat *et al.* 2002). In the recent years, terahertz quantum cascade lasers have made rapid progress. They depend on chirped superlattices (Köhler *et al.* 2002a) and the surface plasmon concept (Köhler *et al.* 2002b) employed for a large optical confinement with low propagation losses, to recently, bound-to-continuum transitions and

extraction of carriers via resonant phonon scattering (Williams *et al.* 2003). The lowest emission frequency is now ∼1.39 THz. More recent progress of QCLs in terahertz range is presented in Subsection 2.5.2. One of considerable advantages of QCLs is the ability to tune the terahertz source to avoid regions with no water 'window'. A relevant experiment has been conducted (Lee *et al.* 2006) with the atmospheric window selected at ∼4.9 THz, to achieve real-time imaging of a dried seed pod at a 25 m stand-off (see the relevant image mode in Section 3.2.3). Another important application of terahertz QCLs is in biomedical imaging; this is well illustrated by the work of Darmo *et al.* (2004) in imaging a rat brain section at 3.4 THz (Section 4.2.2).

Emission of relatively strong continuous-wave terahertz radiation from cuprate superconductors is carried out by Ozyuzer *et al.* (2007). Josephson junctions are stacked within the layered high-temperature superconductors. It is an important step toward filling this 'terahertz gap' from 0.5 THz to 1.4 THz, which the photomixing techniques and QC lasers can barely reach.

2.3 Laser sources

2.3.1 Ti:sapphire-based lasers

Ti:sapphire-based lasers were revolutionized in the 1990s, with the invention of self mode-locking techniques. Currently, it is one of the most important laser sources for generating terahertz radiation (Reid and Fedosejevs 2005). The Ti:sapphire-based lasers are frequently applied for the generation of pulsed terahertz radiation used for experiments conducted in this PhD Thesis.

The principle of Ti:sapphire-based lasers are simply presented as follows. It consists of a rod of Ti:sapphire (titanium-sapphire), which can be pumped by a continuous wave (CW) laser source. This light is focused into the Ti:sapphire rod collinearly with the laser axis through the back of one of the mirrors. Dispersion arises from the variation of the refractive index of the crystal material across the gain bandwidth of the laser. The cavity dispersion output from the crystal material is Q-switched and self-mode-locked with the use of a pair of Kerr lenses (Brabec *et al.* 1992) as a saturable absorber and a Pockel's cell modulator. The schematic diagram regarding the Ti:sapphire pulsed laser oscillator is illustrated in Fig. 2.1(b). The output from the laser oscillator, in Fig. 2.1(a), is taken from the end that is opposite the dispersion-compensating prisms. This alignment of the prism pair creates a longer path for red wavelengths propagating through

the prism material, as compared with the blue, which yields a negative dispersion effect. It is illustrated in Fig. 2.1(c). If the prism separation, lp (defined tip to tip), is sufficiently large, the positive dispersion of the material can be balanced. The prism apex angle is cut to guarantee that at minimum deviation of the center wavelength, the incident angle is the Brewster angle. Fig. 2.1(d) shows the optical path of the output cavity dispersion through the prism pairs. The compressed output pulse is finally achieved from chirped cavity dispersion by the application of prism pairs. The photograph of the Ti:sapphire-based laser is shown in Fig. D.1.

2.3.2 Free-electron lasers

Terahertz free-electron lasers (FELs) feature high intensity, combined with short pulse length, easy tunability and variable pulse structure. These features are essential for biological and medical studies (Grosse 2002) and for the exploration of vibrational and configurational molecular transitions (Kato *et al.* 2000, Xie *et al.* 2002) in terahertz regime.

Unlike femtosecond lasers, i.e. Ti:sapphire-based lasers, where quantum-mechanical characters are employed, the free-electron laser (FEL) (Krishnagopal *et al.* 2004), is a classical laser and converts part of the kinetic energy of the electrons into coherent electromagnetic radiation.

Fig. 2.2 shows a schematic of a terahertz FEL. Two major components of the device are: electron accelerator and undulator. Generally, an undulator, such as in a linac or a synchrotron, is used to produce magnetic fields, which then drive an electron beam from an accelerator to achieve coherent terahertz pulsed radiation. Different from the setup of conventional FELs, the terahertz FELs accelerate electrons directly after they leave the photocathode (Volkov *et al.* 2000). A table-top device of the full terahertz FEL system is an additional advantage as its shorter injector-accelerator combination (Grosse 2002).

NOTE:
This figure is included on page 17 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.1. Ti:sapphire-based lasers. (a) Diagram of a basic self-mode-locked Ti:sapphire oscillator shows the cavity layout. The output pulse is output from the end of the cavity dispersion, and then is tuned by a pair of matching extracavity prisms. (b) The schematics for regenerative amplification in Ti:sapphire oscillator. Ti:sapphire rod, 5 mm or less in length, is doped to absorb about 90% of the incident pump radiation. Two concave focusing mirrors, a pair of Kerr lenses, allow passive mode locking, where M1 is a high reflector and M2 is an output coupler. This pair of Kerr lenses form the optical cavity. When light travels between the two lenses, the energy is accumulated in the interval of the gain medium, which causes population inverse. The TFP indicates a thin-film polarizing beam splitter. The Pockel's cell (PC) is a Q switch and consists of voltage-controlled wave plates. The voltage rotates the plane of polarization by 90°, which results in the light being able to pass the medium. If the laser is Q-switched, a brief burst of light excites population inversion and produces a stronger pulse. (c) Prism pairs are used in the control of dispersion; r and b indicate the relative paths of arbitrary long- and short-wavelength radiation. The incident (Brewster) angle at the prism face is labelled by φ_1. The light is reflected in the plane $p_1 - p_2$ in order to remove the spatial dispersion. (a)-(c): After Reid and Fedosejevs (2005). (d) Illustration of the optical path of the output cavity dispersion through the prism pairs. After Kafka and Baer (1987).

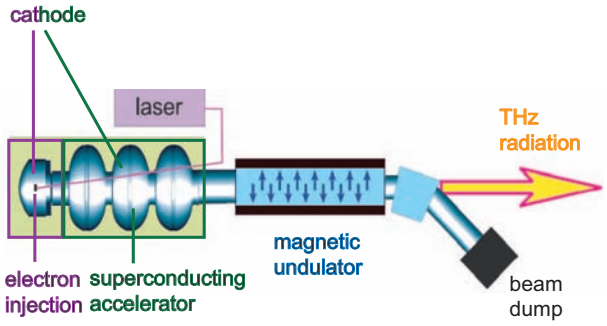


Figure 2.2. A schematic of a THz free-electron laser. The schematic shows a short injector-accelerator combination compared to a stand FEL. After Grosse (2002).

2.3.3 Terahertz quantum cascade laser

Recent advances in nanotechnology have also led to the development of semiconductor-based THz sources: the terahertz quantum cascade laser (Tonouchi 2007). Terahertz radiation is generated when electrons propagate the serial connection of coupled quantum wells. Quantum wells are formed in semiconductors by having a material, like gallium arsenide sandwiched between two layers of a material with a wider bandgap, like AlxGa1-xAs. These structures can be grown by molecular beam epitaxy with choice of the layer thickness, which determine the electron wave functions of the subbands. Unlike conventional bandgap structure of semiconductor lasers, in QCLs, quantum confinement splits the conduction band into a number of distinct subbands (Köhler *et al.* 2002b, Faist *et al.* 1994). The energy spacing of the lasing sub-bands determines the radiation frequency, allowing in principle to produce light at arbitrarily long wavelengths.

Since the first QC laser was demonstrated at the much shorter wavelength of 4 μm (75 THz) at Bell Labs in 1994 (Faist *et al.* 1994), there have been over five types of terahertz QCLs since, due to the different mechanisms of electron energy relaxation processes (Vitiello *et al.* 2006). They are, 'chirped superlattice' (Williams 2007, Köhler *et al.* 2002b), 'interlaced' (Köhler *et al.* 2004), and 'resonant phonon' (Williams *et al.* 2004), 'bound-to-continuum' (Vitiello *et al.* 2006), as well as a recent design of two-colour quantum cascade laser with application of a magnetic field (Scalari *et al.* 2006).

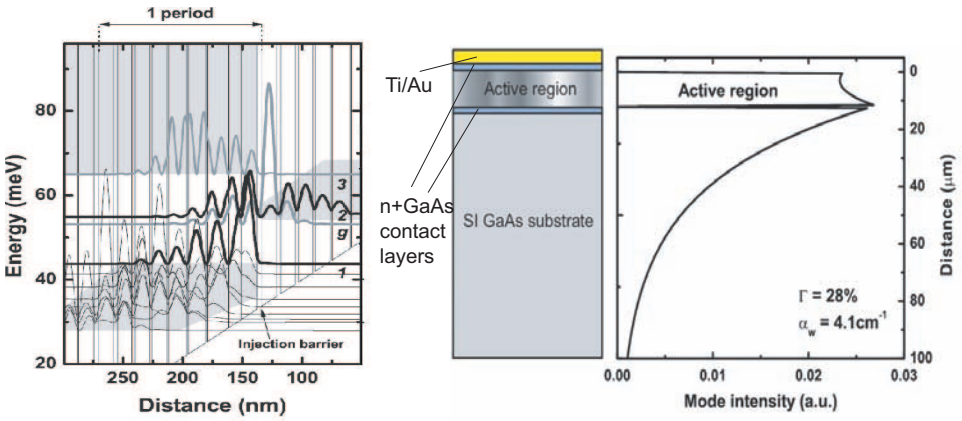
The first THz QC laser with a photon energy less than the semiconductor optical phonon energy was demonstrated at 4.4 THz (equivalent to a wavelength of 67 μm) by Köhler *et al.* (2002b). A chirped superlattice (SL) technique was adopted to design

the QCL active region, with advantage of the long interminiband relaxation time compared to the intraminiband one, and with very large current carrying capabilities and optical powers (Tredicucci *et al.* 1998).

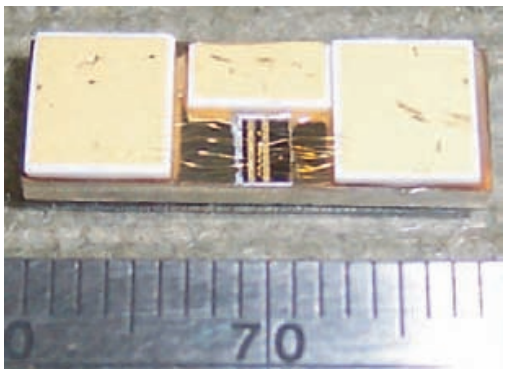
The bound-to-continuum (BTC) QCL has been proved to be successful in producing terahertz radiation with the replacement of the original chirped SL (CSL) design (Köhler *et al.* 2002b). The bound-to-continuum QCL takes place between an isolated upper state and a miniband with a small well adjacent to the injection barrier (Faist *et al.* 2001). The design combines the advantages of the 3 QW and superlattice active regions. It makes it possible to achieve selective injection into the upper state and fast extraction from the lower state. As a result, these designs display improved temperature and power performance compared with the CSL designs (Williams 2007, Alton 2005). A relevant experiment to use a BTC QCL for wavelet based local reconstruction of a 3D polystyrene object is represented in Chapter 13. Fig. 2.3(a) illustrates the design and performance of a bound-to-continuum QCL emitting at 2.9 THz. The laser was made at the University of Cambridge, by Jesse Alton, in collaboration with his colleagues (Alton 2005). The QC structure was grown by molecular beam epitaxy in the $GaAs/Al_{0.15}Ga_{0.85}As$ materials system on semi-insulating GaAs substrates. One period of the conduction band profile for each active region is displayed along with the calculated moduli squared of the most relevant wave functions.

Fig. 2.3(b) illustrates the waveguide design of the two BTC QCLs emitting at 2.9 THz. Both active regions are embedded between identical cladding layers: top contact layer is an 80 nm thick GaAs, and bottom contact layer is a 700 nm thick GaAs. The two contact layers are designed regarding the surface plasmon waveguide. The former has the same mode as a quasi-metallic layer (Köhler *et al.* 2002b) and the latter is completed by a Ti/Au metallization layer on top of the ridge cavity (the top contact layer). Fig. 2.3(c) shows the photography of the finished BTC QCL.

In addition, a THz quantum-cascade laser utilizing alternating photon- and phonon-emitting stages has been developed to achieve efficient extraction of electrons from the lower laser level (Köhler *et al.* 2004). Though the impact of this interlaced photon-phonon cascade laser has been limited, they are particularly notable for achieving very long wavelength operation (Williams 2007). The other major active region type is the Resonant-phonon (RP) scheme. The key design is to use resonant LO-phonon scattering to selectively depopulate the lower radiative level, while maintaining a long upper



(a) (b)



(c)

Figure 2.3. Schematic diagrams of terahertz BTC quantum casecade laser designs. (a) One period of the conduction band profile regarding a 'vertical transition' design of the layer sequence. The upper, and lower state of the laser transition and the injector miniband ground state are labelled 2, 1, and g, respectively. The transition energy, given by the energy difference between 2 and 1, is 12 meV. At alignment the wave functions of the upper and lower state are broadly overlapped. The upper state wave function is concentrated mainly in the two quantum wells adjacent to the injection barrier, reducing the overlap with the lower energy states in the injector miniband, thereby enhancing the non-radiative upper state lifetime. Electrons are injected into the upper state from state g through the injection barrier. The active region consists of 85 identical repeat periods. (b) Schematic representation of completed wafer cross-section (left) and computed optical mode profile (right) along the growth axis. This waveguide is shown at $\lambda = 103$ μm (12 meV) and the yielded waveguide losses (α_w) of 4.1 cm^{-1} and a confinement factor (τ) of 28%. After Alton (2005). (c) Photograph of the BTC QCL. (Figure courtesy of Lynn Gladden, Department of Chemical Engineering, The University of Cambridge, Cambridge, UK)

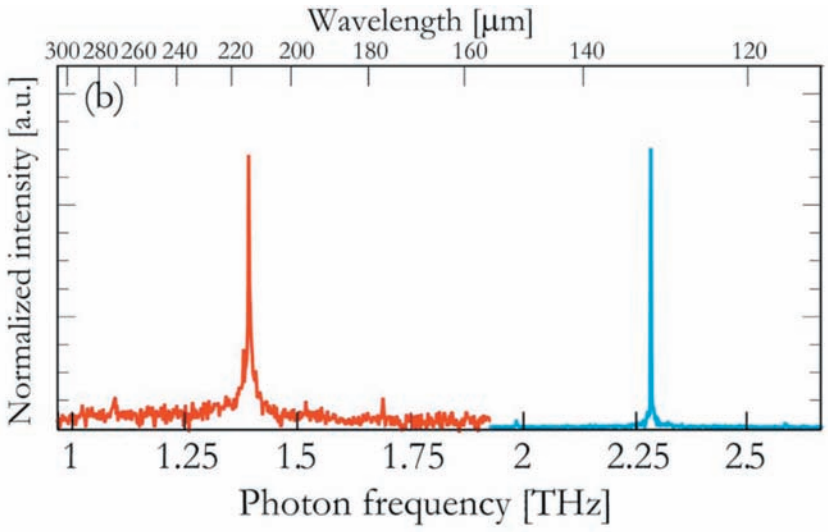


Figure 2.4. Two-colour terahertz quantum cascade laser spectrum. Illustration of the measured spectrum depending on the applied magnetic field. The blue solid line indicates the emission frequency of 2.3-2.36 THz and the red solid red line corresponds to the emitted laser frequency at 1.39 THz. After Scalari *et al.* (2006)

level lifetime. Owing to the lack of a miniband, the RP designs tend to have smaller oscillator strength (Williams *et al.* 2004).

At present, a quantum cascade laser emitting at as low as 1.39 has been demonstrated by Scalari *et al.* (2006). An electrically switchable, two-colour quantum cascade laser emitting at 1.39 THz and 2.3 THz is realized by multi-wavelength operation. A magnetic field is applied perpendicularly to the layers to increase the gain enabling laser action. The structure is based on a large single quantum well and multi-wavelength operation is obtained by selectively injecting carriers via resonant tunneling into one of the excited states of this large quantum well (Sirtori *et al.* 1998b). Fig. 2.4 shows the measured spectrum depending on the applied magnetic field.

2.4 Terahertz semiconductor sources (THz Emitters)

All-optical techniques have been recognized as an alternative approach to produce THz radiation (Davies *et al.* 2002) owing to the difficulties in fabricating solid-state THz sources. Semiconductor surfaces are widely used in conjunction with femtosecond visible/nearinfrared lasers as THz emitters. Bulk electrooptic rectification (difference frequency mixing) and ultrafast charge transport techniques have been exploited. This

Section will review how semiconductor structures can be used to convert ultra-short ($\sim$ 100 fs) near-infrared pulses into THz pulses, and then discuss one THz generation mechanism based on the ultra-fast transport of electrons and holes at semiconductor surfaces.

2.4.1 Bulk electrooptic rectification (optical rectification)

Optical rectification is viewed as one of the main mechanisms to generate terahertz radiation using laser (Tonouchi 2007). When a semiconductor crystal is submerged in a large peak electric field of visible/near-infrared pulse, it induces a second order nonlinear susceptibility ($\chi^{(2)}$) of the crystal. Terahertz radiation can be achieved via exploring the difference frequency of the second harmonic.

Considering a time-dependent polarisation induced in the THz frequency range, it is proportional to the intensity of the incident pump pulse (Davies *et al.* 2002) and the size of the second-order susceptibility. The generated THz frequency ω_{THz} is the difference between the pump pulse frequency ω_1 and the frequency of idling beam ω_2. This result holds when the energy conservation (sum frequency) $\omega_1 = \omega_2 + \omega_{\mathrm{THz}}$ and the phase matching condition $k_1 = k_2 \pm k_{\mathrm{THz}}$ between the induced THz field and the optical fields are satisfied (Schmuttenmaer 2004). The short duration of the incident pulse in the time domain allows a broad bandwidth ($>$ 10 THz) (Ma and Zhang 1993). Since the optical pulses have a bandwidth of a few THz, the difference frequencies fall in the THz range (Schmuttenmaer 2004).

As the ultra-fast pulse generates a polarisation transient $p(t)$ in the EO crystal, according to the Hertzian dipole model, the resultant THz magnitude, $E_{\mathrm{THz}}(t)$, is given by the second time derivative of the polarisation transient, $E_{\mathrm{THz}}(t) \propto \partial^2 p(t)/\partial^2 t$, where the time domain polarisation transient is expressed by $p(t) \propto \chi^{(2)} E^2_{\mathrm{opt}}(t)$, and $E_{\mathrm{opt}}(t)$ indicates the time series of optical pump pulse (Zhang *et al.* 1992). Fig. 2.5(a) illustrates non-resonant optical rectification with an ultrafast laser pulse.

In a material with normal dispersion, the refractive index n increases with the frequency. Since $k = n\omega/c$, it turns out that the k_1-vector is too large, which causes difficulty in satisfying the phase matching condition. An alternative approach to obtain the phase matching condition is by exploiting the k-vectors of the extraordinary and the ordinary axes (Kubler *et al.* 2005). Further research about the nonlinear phase matching for terahertz generation has been conducted (Berger and Sirtori 2004, Sun *et al.* 2007).

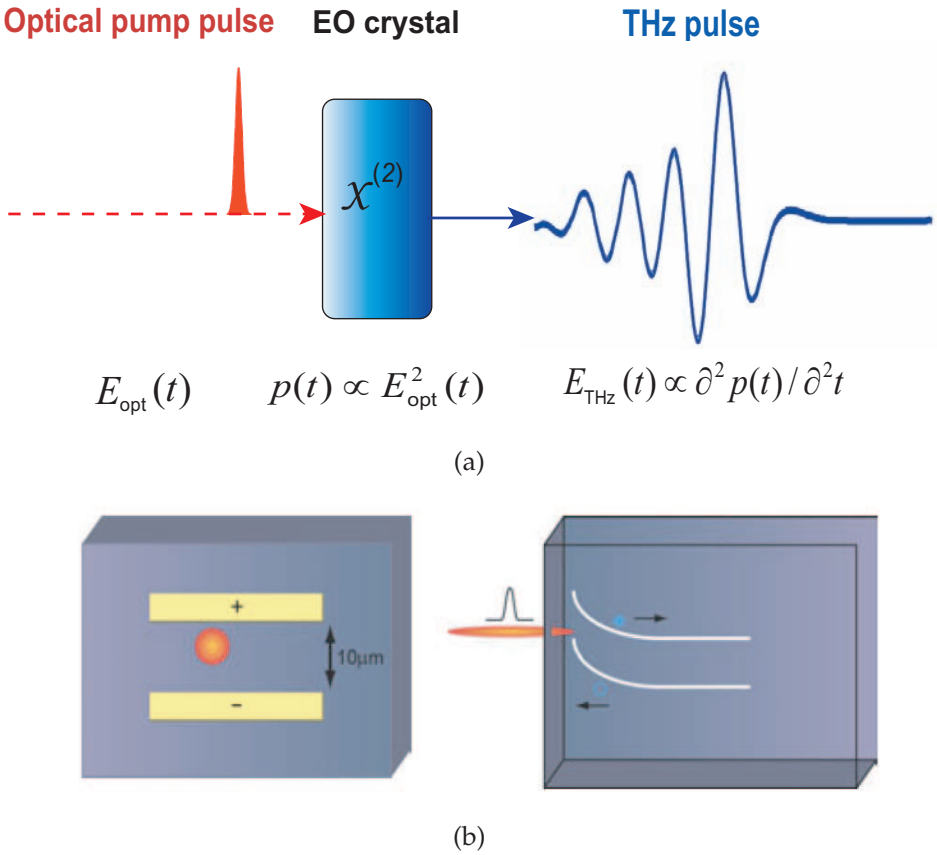


Figure 2.5. Optical rectification and ultra-fast charge transport schemes. (a) Optical Rectification (OR) is a second-order nonlinear effect, whereby an ultra-fast electric field pulse is rectified in a $\chi^{(2)}$ medium, in this case an electrooptic (EO) crystal. The ultra-fast pump pulse induces a transient polarisation, $p(t)$, which in turn emits a T-ray-bandwidth pulse. The time evolution of the T-ray pulse is given by the second time derivative of the polarisation transient, $E_{\mathrm{THz}}(t) \propto \partial^2 p(t)/\partial^2 t$. (b) Illustration of schematic diagrams regarding two THz emitters via applying the ultra-fast charge transport scheme. Photoexcited carriers are generated by a focused visible/near-infrared laser pulse, and then are accelerated and radiate in an electric field. On the left hand side, the field is provided by a lateral antenna structure, on the right hand side, the intrinsic semiconductor surface depletion/accumulation field is exploited. After Davies *et al.* (2002).

To date, this mechanism has been used to generate THz radiation in a number of inorganic semiconducting materials including GaAs, GaSe, ZnTe (Ma and Zhang 1993, Saeta *et al.* 1993, Rice *et al.* 1994), CdTe, DAST, and LiNbO3, although more effective materials are being founded (Davies *et al.* 2002). Zinc Telluride (ZnTe) is regarded as the most popular EO material due to its physical durability and excellent phase matching properties (Rice *et al.* 1994). The organic crystal DAST has a very large EO coefficient (Zhang *et al.* 1992), but is difficult to use experimentally owing to its hygroscopic character. Saturation of OR due to second harmonic generation of the pump beam at high optical high optical fluences has been studied in ZnTe (Sun *et al.* 2000), LiTaO3, LiNbO3 and DAST (Carrig *et al.* 1995).

2.4.2 Ultra-fast charge transport

In an absorbing medium, such as $\langle 110 \rangle$ or $\langle 111 \rangle$ oriented semiconductor, terahertz radiation can be produced by means of 'non-resonant' optical rectification (L. Dakovski and Shan 2005). But if the photon energy of the ultra-fast pulse is higher than the semiconductor bandgap, photons are absorbed, and electron-hole pairs are created close to the surface of the generation crystal. The terahertz radiation is generated primarily as the result of another process mechanism—ultra-fast charge transport. In this process, an appropriate electric field is introduced to accelerate photoexcited electron-hole pairs in semiconductor structures, and the yielding changing dipole leads to generation of a THz pulse.

Typically, suitable surface fields are realized in two ways. On the left hand side of Fig. 2.5 (b), a lateral antenna comprising two electrodes deposited onto a semiconductor surface is shown. A large electric field is applied between the electrodes. The incident laser pulse is focused between the two electrodes, and then creates electro-hole pairs. The free carriers are accelerated by the static field along the field direction and form a transient photocurrent (Zhang and Auston 1992). The emission bandwidth can be modified by appropriate bands structure engineering (Leitenstorfer *et al.* 1999, Davies *et al.* 2002).

A second mechanism is illustrated on the right hand side of Fig. 2.5(b). This was first demonstrated by Zhang *et al.* (1990). A surface electric field is built up when an ultra-fast laser pulse with the photon energy is greater than the semiconductor band gap. The surface depletion field is vertical to the semiconductor surface, which

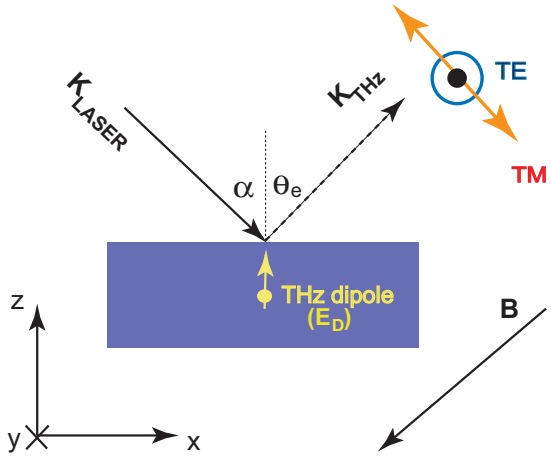


Figure 2.6. Magnetic-field-enhanced generation of T-rays in semiconductor surfaces. An femtosecond laser pulse is incident on the semiconductor surface at angle α, resulting in a depletion field (E_D) vertical to the surface of the semiconductor. The carrier acceleration by the depletion field generates two components: the electric field E_x and E_z of the THz pulse. Both components determine the measured THz signal, which is polarized in the x-z plane. The emitted THz radiation is measured at angle θ_e. The direction of the applied magnetic field is shown by the vector **B**. The emitted THz radiation is collected parallel to the applied magnetic field and perpendicular to the incident beam. The TM mode and the vertical TE mode are labelled by a yellow double arrow line and a blue circle with a black dot inside. After Davies *et al.* (2002).

drives the ultra-fast charge transport. Compared to wide band-gap semiconductors, narrow band-gap semiconductors show an enhanced photo-Dember effect, owing to these characteristics with respect to high electron motilities and large amount of kinetic energy of the photoexcited electrons. The efficient mechanism of the ultra-fast buildup and relaxation of the photo-Dember field results in the generation of THz radiation in narrow band-gap semiconductors (Gu *et al.* 2002). The emitted power and bandwidth for this generation of terahertz radiation is determined by the temperature and properties of the semiconductor crystal, as well as the energy, pulse width and flux of the incident laser pulse (Davies *et al.* 2002).

Recently, there has been enormous interest in this mechanism of the magnetic field enhancement of the terahertz (THz) emission from photoexcited carriers in the surface depletion region of a semiconductor (Izumida *et al.* 1999, Weiss *et al.* 2000, McLaughlin *et al.* 2000). In this case, optical rectification process is negligible owing to the crystal

orientations and the large amount of excess photon energies in the excitation. The analysis of the THz emission process is based on the Drude-Lorentz model (Shan *et al.* 2001, Weiss *et al.* 2000). In the absence of a magnetic field, the emitted THz radiation is completely polarized by transverse-magnetic (TM), caused by the surface depletion field. As the magnetic field is increased, a transverse-electric (TE) component is introduced and the emitted THz radiation becomes elliptically polarized. Both TE and TM power are significantly enhanced in the additional magnetic field. The emission can be measured either by an incoherent bolometeric detection scheme or by coherent electrooptic sampling (Davies *et al.* 2002). Fig. 2.6 illustrates the schematic of the experimental geometry based on the magnitude enhancement scheme.

2.4.3 Terahertz photomixing

Without requesting bulky and expensive femtosecond lasers, photomixing offers an alternative method for generating terahertz radiation. Photomixing can be achieved using two continuous-wave (CW) lasers, and interference produced between two CW lasers (Brown *et al.* 1993, Brown *et al.* 1995) contains different above-bandgap (visible or near-infrared) wavelengths. In these all-photoconductive systems, inexpensive, compact and tunable diode lasers (Gregory *et al.* 2004, Vitiello *et al.* 2006) can be used as the laser sources.

The photomixing techniques based CW-THz emitters and detectors mainly consist of a photoconductor and an antenna. The photoconductor can be viewed as a combination of a photoconductive switch (semiconductor) and a photomixer. The electrodes and photoconductive gap are generically termed the photomixer. The electrodes are designed as the structure of interdigitated fingers coupled to a photoconductive gap. A schematic of a typical CW-THz device is shown in Fig. 2.6(a).

As for a fully optoelectronic detection of the terahertz continuous wave radiation, a component separated from the original laser beam can be used to gate the receiver. A rapid-scan delay line allows collection of the detected terahertz waveforms. A standard 60° 'bow-tie' antenna, shown in the insets of Fig. 2.6(b) and hyper-hemispherical silicon lenses can couple the radiation in and out of the semiconductors (Gregory *et al.* 2004).

NOTE:
This figure is included on page 27 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 2.7. Photomixing schematics. (a) A schematic of a typical CW-THz device. (b) Two insets are the micrographs of the actual structures: a 60° 'bow-tie' antenna structure with a bare photoconductive gap, and the electrode structure linked to the antenna with the design of interdigitated fingers. For both structures the antenna radiation impedance is plotted as a function of frequency. After Gregory *et al.* (2004).

2.5 Terahertz optical sampling techniques

Optical sampling (also called optoelectronic sampling) refers to a class of techniques, which are used to perform time-resolved measurements of terahertz responses (Schmidt-Langhorst and Weber 2005). The optical signals, usually in the form of ultrashort pulses, are used to probe the terahertz electrical field, and ultimately, extract sample information from terahertz responses. In terahertz detection systems, there are four main sampling techniques: free-space electrooptic sampling, antenna detection, synchronous, and asynchronous optical sampling.

2.5.1 Coherent terahertz radiation detection

Coherent terahertz radiation detection techniques involve free-space electrooptic sampling and antenna detection. The electrooptic sampling (EOS) shows good sensitivity and a broad bandwidth, compared to antenna detection. Electro-optic sampling is an optoelectronic technique of optical sampling, realised via exploiting the linear electrooptic effect (also called Pockel's effect). The electric field is exercised by an ultrashort optical pulse, an electrooptic probe beam, and the influence between the electric field and the terahertz radiation occurs during only a short time interval. This effect is usually a change of polarisation, which is turned into a change of optical power by

NOTE:
This figure is included on page 28 of the print copy of the thesis held in the University of Adelaide Library.

Figure 2.8. Sketch of an electrooptic sampling setup. The polarised T-ray electric field induces a birefringence in the detector, according to the $\chi^{(2)}$ coefficient of the specific crystallographic orientation of the crystal (Chen and Zhang 2001). Produced birefringence in turn rotates the polarisation of the probe beam, after the probe beam transmits through the EO detector collinearly with the pump beam (incident T-ray responce), directed by a pellicle beam-splitter. A Wollaston beam splitter, directs the two polarisations to balanced photodiodes. A quarter-wave plate (QWP) is rotated to balance the difference current to zero for zero T-ray field in the EO crystal. After Mickan (2003).

a polariser. A polarising beam splitter, which can be viewed as a second polariser rotated 90° from the first polariser, is used to split and direct the polarisation rotation that the THz field induces on the probe beam and to achieve intensity modulation of the polarisation. This is a typical crossed-polariser detection method (Saleh and Teich 1991, Yariv 1991). A pair of balanced silicon photodiodes is used to analyze the split polarization rotations. The slow variation in the arrival time of the probe pulse, i.e. sequential sampling of a repetitive signal, results in the measurement of a full waveform of a periodic signal. It is observed that electrooptic detection is noise sensitive and there exists a clear trade-off between the noise sensitivity and frequency response. It is determined by the choice of crystal and its thickness (Cai *et al.* 1998). Fig. 2.8 illustrates the sketch of an electrooptic sampling setup.

Photoconductive sampling is an optical sampling technique, which is achieved by using photoconductive switches. It is found that the antenna detection scheme always has a better SNR than EOS, with the chopping frequency at a few kilohertz (Cai *et al.* 1998). Photoconductive detection is similar to photoconductive generation. A short probe laser pulse can close an electrical connection for a very short time. A

NOTE:
This figure is included on page 29 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 2.9. Sketch of a PCA used for photoconductive sampling. This sketch shows a photoconductive dipole antenna used for T-ray detection. The large T-ray spot (shown as a blue circle) biases the electrodes with a free-space electric field. The probe spot is focussed into the biassed region of the antenna substrate. The fs-duration probe switches the electrodes with transient photocarriers, allowing a current to flow, which is detected by an ammeter. The current flow is proportional to the applied T-ray field. After Mickan (2003).

bias electrical field across the antenna leads is generated by the electric field of the THz pulse focused onto the antenna. The photoconductive antenna (PCA) directly detects a photocurrent induced by the incident THz field. Fig. 2.9 is the sketch of a PCA used for photoconductive sampling. The presence of the THz electric field generates current across the antenna leads, which is usually amplified using a low-bandwidth amplifier. This amplified current is the measured parameter which corresponds to the THz field strength. Again, the carriers in the semiconductor substrate have an extremely short lifetime. Thus, the THz electric field strength is only sampled for an extremely short time interval (fs) of the entire electric field waveform, the temporal position of which can be adjusted via an optical delay line (see Fig. 2.10).

2.5.2 Synchronous and asynchronous optical sampling

Synchronous optical sampling is an important implement in a terahertz pump-probe system. In this system, data are recorded in the time domain via different delay times. The varied delay time is achieved by several approaches: (i) traditional scanner with a motorized scanning stage (ii) raster scanning (iii) mechanical scanning devices. This

NOTE:
This figure is included on page 30 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 2.10. Pump-probe delay stage. The delay stage pictured here is the essential element in pump-probe experiments, of which T-ray spectroscopy is one example. The stage is used to sweep the ultra-fast probe pulses over the temporal profile of the free-space T-ray pulse, thereby sampling the entire T-ray waveform. The stage motion is coordinated with the detection electronics via a computer and General Purpose Interface Bus (GPIB) interface. After Mickan (2003).

approach, called synchronous sampling, needs to guarantee correct match between each acquired waveform and the location of the measured object.

Asynchronous optical sampling (ASOPS) is a technique for fast measurements of time-domain spectroscopy (Yasui *et al.* 2005, Bartels *et al.* 2007). The application of this technique has been successfully demonstrated in the field of terahertz time-domain spectroscopy (TDS), i.e. for DNA analysis (Janke *et al.* 2005, Kistner *et al.* 2007) and water vapor detection (Brown *et al.* 2006). This technique adopts two different mode-locked lasers with a slight difference in pulse repetition rates. The two different pulses automatically provide a temporally varying delay. The scan rate is determined by the different pulse repetition rates. Only the low enough scan rate is effective to allows for a sufficient temporal resolution, which is then limited by the detection bandwidth (not by the pulse duration). A high pulse repetition rate helps, as long as a sufficiently large temporal range scanning is available. The large number of the recorded spectra makes possible either to record data for many samples per second (e.g., to obtain position-dependent transmission spectra or even two-dimensional transmission images), or to

average many spectra data from a single sample within a few seconds, aiming to reduce detection noise to a very low level. Fast detection electronic devices allow for recording thousands of transmission spectra, ranging from virtually zero to several terahertz, within one second. Asynchronous sampling avoids mechanical noise of the delay time and position-dependent mode size, but at the cost of requiring two lasers instead of one (Chan *et al.* 2007).

2.6 Chapter summary

This Chapter has broadly described the common techniques and systems for terahertz detection and generation. Several advanced T-ray sources which are primarily used for current popular terahertz research, include (i) pulsed T-rays based on ultra-fast laser sources; (ii) high frequency electronic sources, for their integration into existing electronic technology, to achieve low-power CW operation; (iii) Quantum Cascade Lasers (QCLs), for their small size and tunability, to realise CW operation. The pulsed nature of ultra-fast T-ray systems provides high signal-to-noise ratios, broad bandwidth and low average power, making them ideal tools to study biological and medical materials. However, QCLs as terahertz narrow band laser sources show deeper penetration, due to higher average power, which complements THz pulsed imaging systems. Based on instruments reviewed in this Chapter, the following Chapter reviews T-ray imaging modalities, which underpins THz tomography. Such systems then open up the possibility for imaging and pattern recognition of heterogeneous layers within a target object (Chapter 4).

Chapter 3

Terahertz Imaging Modes

THERE is significant interest in the study of terahertz spectroscopy and imaging particularly dealing with the interaction of terahertz radiation with biological molecules, cells, and tissues (Smye *et al.* 2001). These technologies are capable of revealing spatial and spectral features, such as external and internal structures of a target object, with submillimeter resolution. Since the first illustration of 2D terahertz imaging by Hu and Nuss (2001), terahertz sources and detectors have made much progress, leading to high signal-to-noise ratios and reasonable acquisition rates. Terahertz instruments, especially in biomedical imaging, have the potential to assist biochemists, biologists, medical scientists, and physicians to see their objects of study and to obtain quantitative measurements.

This Chapter mainly provides a taxonomy of various terahertz imaging techniques based on several criteria: operation within versus below the diffraction-limit, pulsed versus continuous imaging, and spectral resolution versus temporal resolution, which are followed by the introduction to a few key types of terahertz time domain spectroscopy (TDS).

3.1 Three fundamental types of terahertz propagation

Terahertz imaging is a relatively new technique and has shown its potential in research and commercial application. Terahertz imaging with high-power THz sources is pulsed scanning and pulse-gated detecting, and the resultant measurement at each pixel position of an image is an entire time-dependent waveform. Therefore, the result from terahertz imaging is a three-dimensional (3D) data set, which then can be mapped to two-dimensional (2D) images (Herrmann *et al.* 2000). Withayachumnankul *et al.* (2007) illustrate such a 3D time-domain array for pulsed terahertz transmission time domain imaging.

3.1.1 Transmission-type terahertz imaging

Terahertz image data contain a wealth of information. The Fourier transform of the time domain waveform in a certain spectral range allows the extraction and display of relevant information of target objects in real time (Mittleman *et al.* 1996), e.g. the calculation of refractive index and absorption coefficients. Another important role of terahertz time domain spectroscopy is to extract phase information; this is achieved by varying the time delay between the THz wave and the probe beam (Mittleman *et al.* 1996). As many materials are transparent to terahertz radiation, it is feasible to measure transmitted responses and acquire spectral information for signal analysis and to produce contrast images.

The time-resolved THz spectrometer used in most of the studies presented in this Thesis utilizes a short coherence length infrared source (centered at around 800 nm) to generate a sub-100 femtosecond duration pulse train with repetition frequency of around 80 MHz. Each infrared pulse, is split into separate pump and probe beams. The pump beam is used to excite an optical rectification crystal, which acts as a T-ray emitter, and the T-rays produced (duration around 200 fs) are collimated and focused onto a sample by a pair of parabolic mirrors. The T-rays emerging from the sample are re-collimated by another pair of mirrors, before being combined with the probe beam in a T-ray detector crystal. As a result, the modification by the sample T-ray and the probe beams propagates through the THz detector crystal co-linearly. The pump beam, which is

NOTE:
This figure is included on page 35 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.1. Transmission mode THz imaging. A schematic experimental setup for electrooptic transmission terahertz imaging with ZnTe as EO generation and detection, illuminated by a femtosecond laser. The inset on the bottom is the plot of a time domain THz waveform before and after passing through pork-fat tissue, indicated by solid and dotted curves, respectively. After Han *et al.* (2000).

also transmitted through a chopper, travels through an optical delay stage that is modulated accordingly, so that the pump and probe beams arrive at the detector in a time-coincident manner. The electrooptic detector crystal produces an output that is proportional to the birefringence observed from the interaction of the THz pulse with the time-coincident infrared pulse replica within the crystal. This output is proportional to the T-ray response of the sample and this signal is measured with the use of an optical photodetection scheme. A lock-in amplifier is also used to demodulate the signal, and this avoids $1/f$ (flicker) noise problems that are present in this detector-limited measurement scheme. Terahertz pulsed imaging (TPI) is achieved by performing a 2D raster scan after translating the sample in both the x and y direction, while keeping it at the focal plane of the parabolic mirrors. A typical setup (Ferguson *et al.* 2003) is shown in Fig. 3.1.

The inset on the bottom in Fig. 3.1 is the plot of a time domain terahertz waveform before and after transmitting through 0.5 mm pork-fat tissue. The dotted curve of the target measurement reveals a phase shift, broadening, and amplitude attenuation.

The phase shift results from the average refraction index of the sample, whereas the broadening is caused by dispersion and frequency-dependent attenuation.

3.1.2 Reflection-type terahertz imaging

On the one hand, terahertz waves can penetrate most dry, nonmetallic and nonpolar objects, such as teeth, foam, leaves, some fatty tissues, while on the other hand, certain materials, such as metals, are completely opaque to terahertz radiation (Mittleman *et al.* 1996). A THz pulsed imaging (TPI) system operating in reflection mode might be a better choice for a number of practical applications. For example, transmission terahertz spectral measurements show that large sucrose particle samples with an average thickness of approximately 0.5 mm are actually 'opaque' to terahertz radiation (Mittleman *et al.* 1996). In contrast, as demonstrated by Shen *et al.* (2005), a TPI system operating in reflection mode has no such limitation and can still be used to obtain reliable terahertz spectra.

Fig. 3.2 is a schematic experimental setup for photoconductive reflective terahertz imaging system using a pump-probe configuration. In this diagram, the path lengths between the pump beam and probe beam are adjusted by an optical delay stage to achieve equal length of arms, therefore distance measurements are possible within the coherence length of the THz radiation. Like optical coherent tomography (Huang *et al.* 1991, Tearney *et al.* 1995), the correlation signal can then be measured to provide information about the reflectivity of the sample. The ultimate spatial resolution achieved in depth (i.e. the propagation direction of the reflected beam) is limited by the coherence length of a terahertz pulse, which is determined by the pulse duration, and is to the order of several hundred femtoseconds. The *x*-*y* stage is to achieve terahertz measurements pixel by pixel. The transient electric field measured by the reflection-mode TPI system using a photoconductive emitter and detector shows excellent signal-to-noise ratio (S/N) and high dynamic range (Pickwell and Wallace 2006).

The inset on the left hand side of Fig. 3.2 illustrates a time domain terahertz waveform after reflection from a film-coated tablet. The dominant peak *a* corresponds to the air-tablet surface interface. The negative peak *b* represents the interface between film coat and tablet core. Since each peak in the terahertz waveform corresponds to a different interface within the sample, by searching for peaks, the time delay of the

Figure 3.2. Reflection mode THz imaging. A schematic experimental setup for photoconductive reflective terahertz imaging, illuminated by a femtosecond laser. After Woodward *et al.* (2003). The inset on the left hand side is the plots of a time domain terahertz (THz) waveform after reflection from a film-coated tablet. After Zeitler *et al.* (2007b).

terahertz pulse can be calculated and the thickness information of the coating layers can be determined. Further forms of T-ray reflection imaging are illustrated by (Mittleman *et al.* 1996, Woodward *et al.* 2002, Pearce *et al.* 2005).

Though terahertz reflection imaging permits us to achieve depth resolution of target objects, the difficulty lies in the achievement of sufficiently accurate phase measurements. The main problem is that the small misalignment with precision lower than 1 μm between the surfaces of the sample and a reference mirror can cause a considerable error (Jeon and Grischkowsky 1998, Nashima *et al.* 2001). Therefore, terahertz ellipsometry imaging is proposed by Nagashima and Hangyo (2001) to measure the frequency dependence of the complex refractive index of samples without a reference measurement. Although the S/N ratio of ellipsometry-type THz spectra is insufficient in the high frequency region, illustrated by Hangyo *et al.* (2002), this can be improved by changing the incident angle to be near the Brewster angle.

3.2 Terahertz imaging within diffraction-limit

One of the clear limitations of all of the terahertz imaging techniques is the spatial resolution (Chan *et al.* 2007, Johnston 2007). Spatial resolution, which is also called angular resolution, describes the ability of an optical imaging device to measure the angular separation at each point of an object in this context. The spatial resolution of optical devices is ultimately limited by diffraction.

The applications of terahertz imaging within the diffraction-limit, represented in this Chapter, cover topics ranging from time-of-flight imaging to tomography with pulsed terahertz radiation; from an all-optoelectronic continuous wave THz imaging to imaging using a THz quantum cascade laser. All the imaging setups in this Thesis are based on terahertz time domain spectroscopy techniques.

3.2.1 Terahertz time-of-flight imaging

Temporal resolution is a term used to describe measurement precision with respect to time. Often there exists a tradeoff between temporal resolution and spatial resolution. Time-of-flight imaging enables the measurement of depth information. Time-of-flight imaging in the terahertz range has been widely used (Mittleman *et al.* 1999, Han *et al.* 2000, Zeitler *et al.* 2007b, Pradarutti *et al.* 2007).

Time-of-flight imaging in biomedicine is well demonstrated by Han and Zhang (2001) using a transmission-type TPI with onion cells as a sample. The THz image is achieved by plotting the transillumination amplitude of the THz pulse at a fixed time delay between the THz and probe pulse, which is shown in Fig. 3.3. Thus, the contrast in the image is attributed primarily to differences in the water content of the cells and the intercellular regions.

One disadvantage of the ultrashort pulse technique is that it takes usually several hours to acquire the data in three dimensions (two lateral dimensions and one time dimension) for image scanning. A fast scanning method for image acquisition is introduced by Pradarutti *et al.* (2007), which depends on the refractive index of a sample. The contrast is as good as the full time delay method and better than with amplitude scanning, with reduced scan time. More techniques regarding terahertz time-of-flight imaging are being introduced and improved to prompt the advancement in biomedicine imaging (Pradarutti *et al.* 2007).

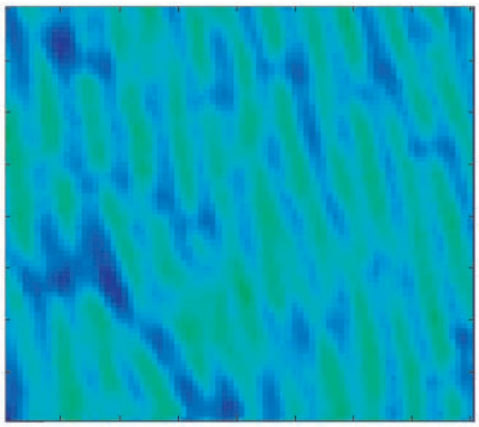


Figure 3.3. Time-of-flight imaging. High resolution of THz image of onion cells is illustrated using TPI. After Han and Zhang (2001).

3.2.2 Tomography with pulsed terahertz radiation

Different from terahertz time-of-flight imaging, the tomographic reconstruction modality has its own source-detector configuration and image reconstruction algorithms, appropriate to the nature of the problem. It is suggested that the number of transmitters and/or detectors is a criteria to distinguish tomographic reconstruction from the time-of-flight images (Chan *et al.* 2007).

The time-of-flight measurements consist of a single transmitter and a single receiver, both at fixed locations. As tomographic reconstructions view objects from multiple angles, multiple illumination sources or multiple signal detectors are needed for the measurements. But normally, a single transmitter-receiver pair, and a hardware configuration for a series of measurements at multiple angles are employed to achieve the description of an image (Ferguson *et al.* 2002b, Nguyen *et al.* 2006, Chan *et al.* 2007).

The first terahertz tomographic measurement was demonstrated by Ruffin *et al.* (2000) using fiber-coupled techniques (Rudd *et al.* 2000). The shape of target objects is recovered by reversing scattered fields mathematically. Single-cycle THz pulses are generated and detected for the object measurements according to the research carried out by Cheville and Grischkowsky (1995).

Two steps are involved for the basic time-reversal reconstruction. (i) The time domain Fresnel diffraction formula is applied for the diffraction of electromagnetic pulses from given terahertz input field; (ii) the reconstructed diffracted field then can be recovered as an input field from time-reversal symmetry of Maxwell's equations. An example

of reconstruction using 2-D apertures is accomplished via rotating the object about its central normal axis and then back-propagating the measured fields using the Kirchhoff diffraction integral. Fig. 3.4(a) illustrates the diffracted field at the off-axis position P_l, which can be mathematically time-reversed and used as an input field to reconstruct the field at the object's position P_o. Fig. 3.4(b) is the 2-D experimental setup. Diffracted electric fields are measured at one off-axis position as the object (spiral) is rotated about the z-axis (Ruffin *et al.* 2002). The rotation angle of the target around the z-axis is indicated by θ_d.

NOTE:
This figure is included on page 40 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.4. Single-cycle THz tomography. (a) Illustration of the diffracted field at the off-axis position P_l, which can be mathematically time-reversed and used as an input field to reconstruct the field at the object's position P_o. (b) Illustration of the 2-D experimental setup with θ_d of 12°. The dashed circle represents the measurement points with θ_d of 12° on the spherical surface. After Ruffin *et al.* (2002).

The sparrow criterion (Sparrow 1916) is applied to derive the resolution resulting from time-reversal reconstruction. It is attractive due to its low computational overhead. Two peaks of measured waveforms are resolved if there is a clear local minimum between the principal peaks of the two waveforms. This resolution criterion allows the high temporal resolution of THz-TDS systems to be leveraged to derive the spatial resolution. The resolution is given by the spatial separation of two points on the object plane that give rise to THz pulses with an observable timing difference at the detector. A star target is employed for the reconstruction experiment. It turns out that there is

NOTE:
This figure is included on page 41 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 3.5. Terahertz multistatic reflection imaging. Illustration of a schematic of a common shot experimental arrangement, emulated by a terahertz (THz) system. After Dorney *et al.* (2002).

an obvious reduction in the number of measured pixels for recovery of the target object compared to a conventional scanned image. It results in a considerable saving in acquisition time and demonstrates the effectiveness of this technique.

At roughly the same time, Dorney *et al.* (2002) realises tomographic reconstruction for geophysical applications. Like time-reversal reconstruction, a two dimensional Kirchhoff migration (Dobrin and Savit 1988, Scales 1995) process is employed to the solve the inverse problem. The imaging algorithm is very simple to implement and is extremely robust against losses due to scattering or absorption of the propagating wave, because only the time-of-flight, rather than the amplitude of the measured wave, is used for image reconstruction.

The task of Kirchhoff migration is to transform this data set, containing the positions of the transmitter and the receivers along with the delay times of the reflected pulses, into a useful image of the subsurface. Fig. 3.5 is a schematic of a common shot experimental arrangement, which can be emulated by a terahertz (THz) system. The horizontal surface is the x-axis and the depth is the z-axis, and the goal is to transform data in the (x, t) domain into the (x, z) domain.

T-ray holography is an extension of recent work in THz imaging using time-reversal of the Fresnel-Kirchhoff equation (Scales 1995). Terahertz holographic reconstruction has the advantage of utilising the multiple scattering features of terahertz radiation. Both the transmitted and reflected waves can be employed for image reconstruction. Compared to conventional continuous wave holography, THz holographic techniques,

in addition to the phase and amplitude information at each frequency, also contain the temporal information that may be used to separate scattered waves with different scattering orders. A method for studying such local spectra corresponding to each separated scattered wave is to truncate the wave in the region of interest and perform the windowed Fourier transformation (Kaiser 1994, Carin *et al.* 1997) with the Born series used for a reconstruction of an image.

An example of combining holographic techniques and terahertz radiation, with a target containing two 3.5 mm thick polyethylene plastic sheets, is illustrated by Wang *et al.* (2004a) and Wang *et al.* (2004b). A schematic of this experimental arrangement is shown in Fig. 3.6(a). The two sheets are labelled S1 and S2, each having triple-hole patterns.

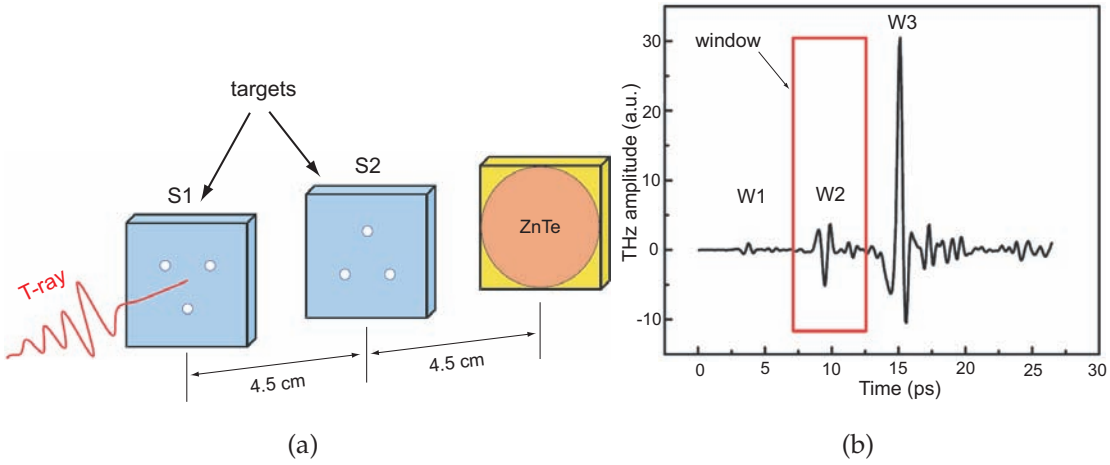


Figure 3.6. Terahertz holography. Illustration of the experimental arrangement. The distances from S1 and S2 to the ZnTe sensor are 4.5 cm and 9 cm, respectively. The interval between the holes on each sample (6 mm) is much larger than the peak value of the THz beam (0.3 mm), and the hole diameters are 1.8 mm. The THz waveform is measured at the centre pixel, and the corresponding plot is shown in Fig. 9(b). It consists of three distinct pulses, W1, W2 and W3, and the start time of each pulse is at around 3.4 ps, 9.2 ps, and 15 ps, respectively. This waveform information is valuable for studying the timing of these pulses at each pixel of the sensor. (b) Plots of the THz waveform measured at the centre pixel. After Wang *et al.* (2004a) and Wang *et al.* (2004b).

Fig. 3.7 illustrates the holographic reconstruction using terahertz radiation. It shows excellent correspondence with the target geometry (shown in (a) and (b)), although the holes in the far plane S2 are slightly blurred. A similar principle can be applied to

achieve images with more point scatterers on more than two target planes (Wang *et al.* 2004a, Wang *et al.* 2004b).

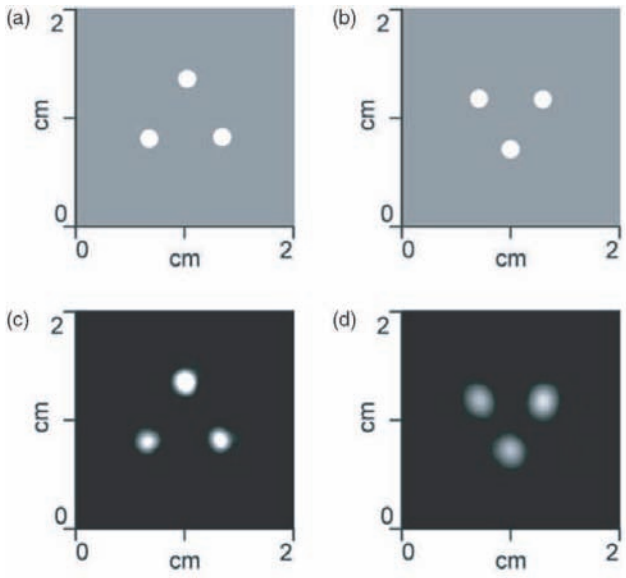


Figure 3.7. Holographic imaging reconstruction. Schematic of the target sample planes and their reconstructed holograms. Schematic of sample S1 (a) and S2 (b) reconstruct hologram of S1 (c) and S2 (d). The distances from the ZnTe sensor to two samples S1 and S2 are 4.6 cm and 9.3 cm, respectively. After Wang *et al.* (2004a) and Wang *et al.* (2004b).

A further development in holographic techniques has been made recently. Compact imaging systems, which rely on viewing an electrooptic crystal using a charge-coupled device (CCD) camera have been achieved. This system has the capability to capture two- or three-dimensional images through holographic techniques, as illustrated by Crawley *et al.* (2006).

The tomographic reconstructions mentioned above are realised with application of the Kirchhoff propagation equation. Terahertz computed tomography in reflection mode is also reported according to the work performed by Pearce *et al.* (2005). Compared to the first demonstration of reflection mode of terahertz tomography, with only a single viewing angle and the Fresnel coefficients for object estimation (Mittleman *et al.* 1997, Mittleman *et al.* 1999), terahertz wide aperture reflection tomography (WART) is capable of detecting objects located behind strong reflectors using multiple viewing angles. The principle of THz reflection tomography is the same as transmission THz computed tomography, first illustrated by Ferguson *et al.* (2002b) (see Chapter 10), using techniques borrowed from X-ray imaging. The reflected waves from target objects

can be viewed as parallel projections through the cross-section, and therefore computed tomography (CT) reconstruction algorithms such as the filtered back projection (FBP) algorithm, can be applied to retrieve an edge map of the object's cross-section. In the terahertz imaging community, more methods have been explored to achieve the analogs of each of these imaging configurations. The study of tomographic techniques using terahertz radiation has been discussed in a recent review (Chan *et al.* 2007). Computed tomographic reconstruction using terahertz radiation is an important part of this Thesis' scope.

An another application of THz systems is in performing radar cross-section (RCS) measurements of scale models of military vehicles and aircraft as an inexpensive alternative to operational trials (Cheville and Grischkowsky 1995, Cheville *et al.* 1997), which was first realised by Pearce *et al.* (2005).

Fig. 3.8(a) is a schematic diagram of the THz impulse range. The receiver is fixed as the target is rotated on a center axis. For the synthetic aperture imaging, the receiver module measures the time dependent scattered field at various angles with respect to the target space. The impulse range configuration shown in this figure collects two dimensional array data (electric field as a function of time and angle) and images two-dimensional target spaces in the x-z plane.

Fig. 3.8(b) shows a high contrast photograph of a metal model, with a scale indicating length. Despite the image being created using a relatively low number of scans, the superstructure with the antenna masts and the side of the ship are clearly observed in the image. One notable aspect of THz impulse ranging is that while the lateral resolution is typically limited by the Rayleigh criterion to the order of the peak THz wavelength, the depth resolution is dependent on the spectral bandwidth of the THz pulses. For a typical pulse with a rise time of $\Delta t = 0.8$ ps the range resolution is $\Delta t/2c = 0.12$ mm, which is almost an order of magnitude smaller than the peak wavelength.

3.2.3 Terahertz continuous-wave imaging

The technique of using continuous-wave radiation for terahertz nondestructive testing has been exploited for several decades (Karpowicz *et al.* 2005b). In recent years, with the advancements in semiconductor technology, both compact and simple optic hardware systems for terahertz CW imaging are possible. Such a low-cost, compact system has image capture rates comparable with those from state-of-the-art pulsed

NOTE:
This figure is included on page 45 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.8. Synthetic aperture radar algorithms using T-rays. (a) Schematic diagram of the THz impulse range. The THz beam on the target is a 15 mm ($1/e$) wide Gaussian THz wave. The target object is viewed as a superposition of point scatterers made at each particular value of x_0. The measured signal of the target object, at the increased stage rotation angle $\Delta\theta$, is a superposition of the field from each point scatterer at a fixed value of x_0. The 2D target image is a superposition of point scatterers made at each particular value of x_0. To reconstruct the image of the target, after conducting Fourier transforms, the complex amplitude is multiplied by a distance dependent phase shift ϕ, for each particular position x_0. The data are then inverse transformed into the time domain, yielding a plot of scattering amplitude as a function of x_0 and $t = z/c$. (b) High contrast photograph of a 1:2400 destroyer model above a THz impulse SAR image is taken with 20° angular range and 1° resolution. After Pearce *et al.* (2005).

THz systems (Gregory *et al.* 2004). There are a series of methods involved to generate continuous wave terahertz radiation for imaging.

Siebert *et al.* (2002) report an all-optoelectronic THz imaging system based on photomixing, which is used for imaging a thin-cut canine basal cell tumour. It is an encouraging technique, especially applied to imaging biomedical samples, though it is not cost effective due to its use of two expensive Ti:Sapphire lasers.

Fig. 3.9 shows the layout of this system. The optical system is similar to pulsed terahertz transmission measurement systems. Instead of one laser, a dual-colour CW Ti:sapphire laser is used (Siebe *et al.* 1999) to generate two laser beams. A beam-splitter

NOTE:
This figure is included on page 46 of the print copy of the thesis held in the University of Adelaide Library.

Figure 3.9. All-optoelectronic continuous wave THz imaging. Schematic of all-optoelectronic continuous wave THz imaging. One optical beam is guided via a computer-controlled optical delay line to the emitter antenna, and the other is used to gate the receiver antenna. The two antennas are for photomixing. Lock-in detection of the THz signal is adopted to avoid detrimental feedback into the two-colour laser source instead of a mechanical chopper. After being coupled between the emitter and a Si substrate lens, the THz radiation is collimated and focused onto sample by off-axis paraboloidal mirrors. The detector antenna records the terahertz responses in transmission mode. A biomedical sample is mounted on a computer controlled x-y stage. The relative phase of the THz signal is varied by translating the optical delay line with variety over two THz periods. A reference scan of the time delay is taken without the sample during a scan of a row of pixels to compensate for any differences in the phase of the signal due to the change of the scanning direction. The inset shows a logarithmic power transmission image taken with the CW system at 1 THz with an object size of 32 mm $\times$ 24 mm $\times$ 3 mm. The numbers from labels 1 to 4: skin, fat, tumour, and connective tissue. The image consists of 11,248 pixels, which takes a 39 minute acquisition time. The image noise for the relative power transmission turned out to be 13:1, corresponding to 26:1 for the field amplitude. After Siebert *et al.* (2002).

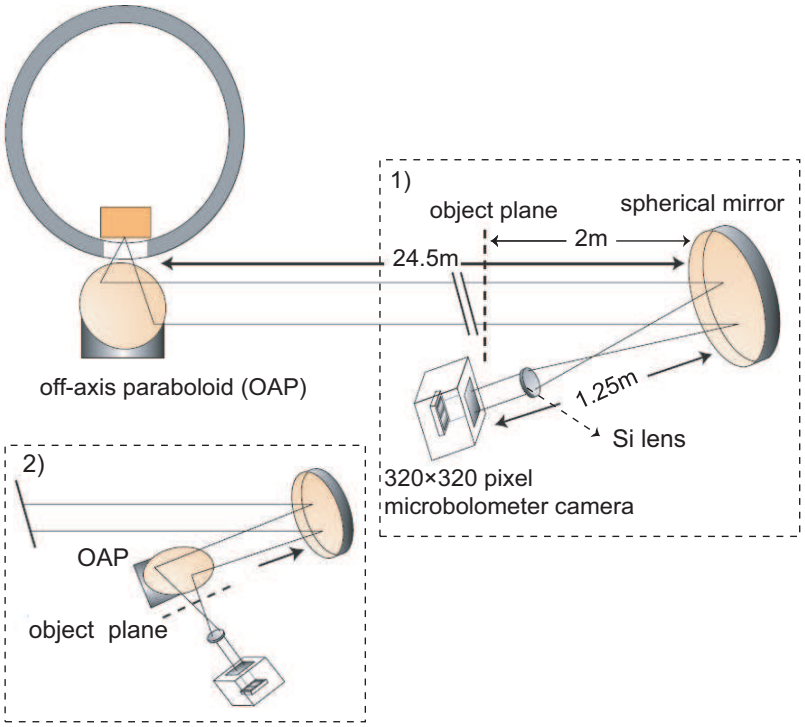
cube combines the two laser beams and then produces a pump-probe configuration. The intensity of the pump and probe beams is modulated at the difference frequency of the two laser beams from 0 THz to 10 THz. Two photomixing antennas are used as emitter and detector of terahertz CW radiation. This system produces high output power and broad detection bandwidth with the growth of temperatures. It is found that the extremely long coherence length of the two optical single-mode frequencies allows the shortest possible imaging data acquisition time.

The CW imaging system is validated by a biomedical sample, a wax-mounted slice through a canine skin tumour. A logarithmic power transmission image taken with the CW system at 1 THz is illustrated in the inset of Fig. 3.9. The resultant experiment is comparable in image capture rate and image noise to pulsed THz imaging systems (Siebert *et al.* 2002).

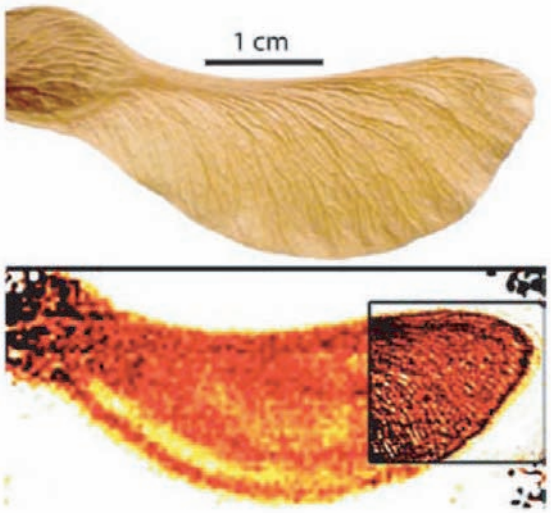
However, there exists a drawback when solid-state electronic devices are used to generate terahertz radiation. The output power rolls off at high frequencies due to transit-time and resistance-capacitance effects (Siegel 2002, Woolard *et al.* 2005). The lack of appropriate materials with sufficiently small band gaps limits photonic approaches to direct terahertz generation. Since the lack of appropriate interband materials, semiconductor terahertz lasers, in the past, relied on more exotic gain mechanisms.

Terahertz quantum cascade lasers (QCLs) are appealing. Their intersubband-transition nature implies that any desired frequency can be achieved by band gap and waveguide engineering. Terahertz radiation, over a continuous frequency range from 1.59 THz to 5 THz, has been demonstrated (Williams 2007).

The special importance of terahertz narrow band sources lies in their ability to select the source wavelength for optimum terahertz imaging capability. A considerable advantage of this QCL is the ability to minimize the atmospheric attenuation via tuning the terahertz source frequencies within a water absorption window. Lee *et al.* (2006) demonstrate a terahertz quantum cascade laser for real-time imaging in transmission mode at a standoff distance of 25 meters. Lasing frequency at $\sim$ 4.9 THz is selected for optimum transmission. It is observed that in the 25 m stand-off imaging experiment, relative transmitted power is increased from 0.32% to 5.6%, if the atmospheric attenuation coefficient is decreased from 1 dB/m to 0.5 dB/m. Sensitivity of QCL imaging could be improved by optimizing the microbolometer absorber for terahertz (about 10 dB improvement), or by moving to cryogenic direct detectors, such as quantum-well infrared photodetectors (Luo *et al.* 2005).



(a)



(b)

Figure 3.10. THz QCL imaging at a standoff distance. (a) Schematic of experimental setup for imaging over a distance of 25.75 m, where a QC laser mounted in a pulse-tube cryocooler is used as the illumination source, and a room-temperature microbolometer focal-plane array camera is used to perform imaging in two possible configurations, (1) and (2). Also shown is a seed pod visible image, (b)-top, and terahertz image, (b)-bottom, taken with a 1 s integration time in an imaging configuration (1). The inset of (b) shows the terahertz image from configuration (2). After Lee *et al.* (2006).

3.3 Terahertz imaging below the diffraction-limit

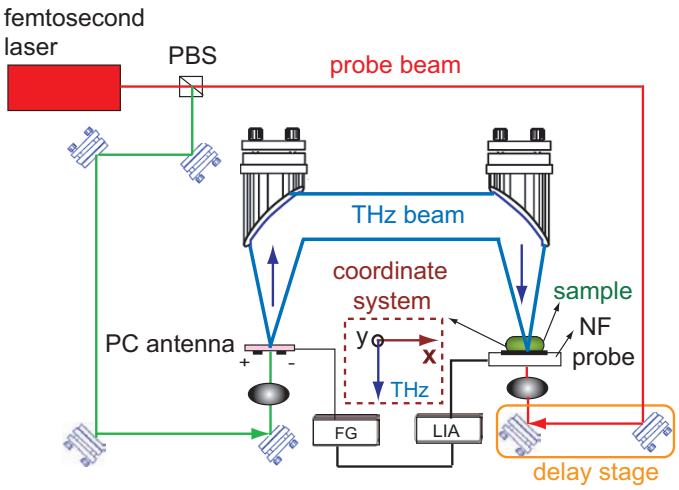
The size of smallest object that a lens can resolve is close to the wavelength, λ. Moreover, the smallest resolvable features in an image are not generally smaller than the wavelength. In order to achieve the image of smaller objects that cannot be resolved in detail, it is needed to break diffraction limit to push the resolving power below the wavelength limit (Chan *et al.* 2007).

A limitation of THz imaging is poor spatial resolution due to the relatively long THz wavelength. Terahertz near-field scanning can significantly improve the spatial resolution. Various methods regarding terahertz near-field imaging have been explored to improve the resolution to a few tens of microns (Mitrofanov *et al.* 1998, Mitrofanov *et al.* 2000, Chen *et al.* 2000). A dynamic aperture technique, used for the terahertz near-field imaging, is capable of improving resolution to a subwavelength range ($\lambda/10$). However, since the spatial resolution of the dynamic aperture is determined by the diameter of the photocarrier layer, it is easily damaged by laser induced heating (Chen and Zhang 2001).

An alternative approach is represented by Federici *et al.* (2002), where a micro machined near-field probe is fabricated by Lee *et al.* (2001). The near-field probe is coupled through an aperture as small as $\lambda/300$. An efficient design of the probe allows detection of terahertz electric field. The THz near-field imaging set-up is illustrated in Fig. 3.11(a). It is similar to photoconductive terahertz pulsed imaging, in addition to the important application of the probe design.

The probe structure is illustrated in Fig. 3.9(b). It is designed with an entrance subwavelength aperture of size d (5-50 μm) on a thinned GaAs layer and a PC antenna detector embedded between a GaAs and a sapphire substrate. A high refractive index material is inserted in the space behind the aperture that reduces the effective wavelength. The sapphire substrate supports the structure and allows the optical gating pulses access the antenna from the substrate side.

The near-field probes based on coaxial transmission lines do not exhibit a cut-off frequency, compared to near-field probes that depend on small apertures that produce THz transmission. The spatial resolution of the near-field probe is defined by the aperture size. It is found that the spatial resolution of larger apertures scales with the aperture size and is independent of wavelength (Mitrofanov *et al.* 2001b).



(a)

NOTE:
This figure is included on page 50 of the print copy of the thesis held in the University of Adelaide Library.

(b)

Figure 3.11. Near-field (NF) imaging. (a) A THz near-field imaging set-up. A mode-locked Ti-sapphire laser beam is split to generate a pump beam (green line) and a probe beam (red line). The pump beam excites a photoconducting (PC) switch. Generated THz pulses (blue lines) are directed by two off-axis parabolic mirrors and focused on the object through an aperture and a transparent substrate. The near-field probe is placed behind the sample and nearly contacts with it. A PC antenna is integrated into the near-field probe to detect the terahertz radiation, gated by probe beams. A lock-in amplifier is used to improve the terahertz response. An automated x-y translation stage scans an object perpendicular to the optical axis. A variable time delay stage allows time-domain sampling of the THz pulse. The image can be constructed using the detected THz signals either at a fixed time delay or at each position of the entire time domain wave form. (b) Illustration of the probe structure, being coupled through an aperture as small as $\lambda/300$. After Federici *et al.* (2002).

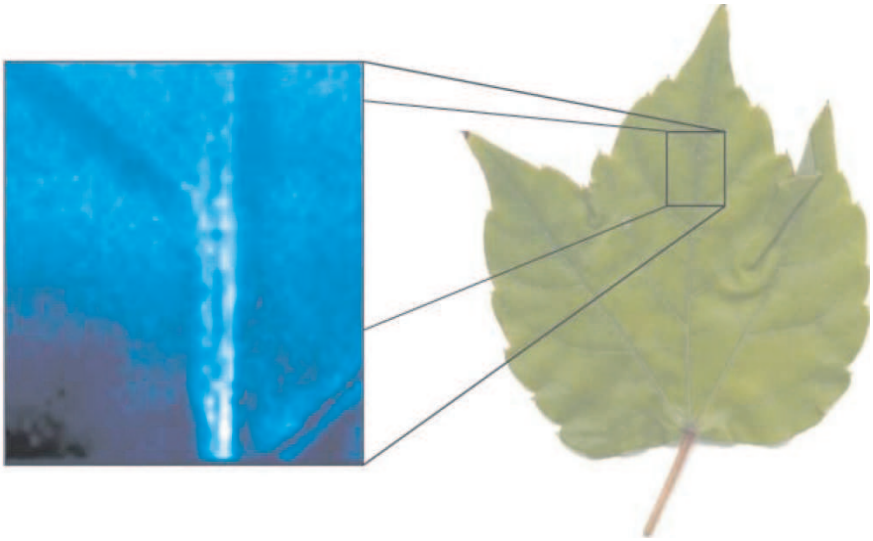


Figure 3.12. THz near-field imaging employing synchrotron radiation. A part of a freshly cut parthenocissus leaf is imaged in transmission. After Schade *et al.* (2004).

As the probe aperture alters the detected waveform, in the near-field, the target object is not imaged directly based on the measurement of the instantaneous terahertz electric field scattered by the object. In principle, the original waveform can be extracted if the transfer function of the aperture is known. Finite-difference time-domain numerical methods (Mitrofanov *et al.* 2001a, Bromage *et al.* 1998, Taflove and Hagness 2001) can be used to analyse THz near-field image formation.

A spatial resolution of 7 μm has been achieved by using a 5 μm aperture probe ($L =$ 4 μm, see Fig. 3.9(b)) (Federici *et al.* 2002). In order to maintain the best THz spatial resolution, in the near-field, with under 7 μm spatial resolution thin sample sections with less then 10 μm thick are used. Measurement of the THz optical properties or spectra of isolated biomolecules (e.g., DNA) are examples of applications of THz near-field imaging (Brucherseifer *et al.* 2000).

One more THz near-field imaging technique based on the broadband highly brilliant coherent synchrotron radiation (CSR) has been reported by Schade *et al.* (2004) at the BESSY synchrotron. This near-field imaging technique enables spectroscopic mapping of strongly absorbing materials at spatial resolution well below the diffraction limit.

Fig. 3.12 shows a transmission image of a part of a freshly cut parthenocissus leaf. In this THz near-field image the inner structure of the veins is revealed, which is formed by spectrally broad water absorption and possible scattering at the structural boundaries. The contrast between the veins and leaf is shown via the 2D mapping. Since it

is sensitive to the water concentration of the sample, THz scanning near-field imaging (SNIM) enables studies of hydration dynamics with high spatial resolution.

3.4 Chapter summary

This Chapter gives an overview of basic terahertz imaging modes for terahertz imaging systems, putting into the context some T-ray experiments conducted by international THz groups. The image modes form a practical platform for image analysis with application of detecting a wide variety of biosamples, for the aims of security screening, quality control and biomedical diagnosis, which are represented in the next Chapter.

Chapter 4

Terahertz Imaging Analysis

ADVANCES in THz technology have propelled terahertz (THz) radiation to be an important part of the electromagnetic spectroscopic region. The THz frequency range excites large amplitude vibrational modes of molecules, and allows researchers to probe weak interactions (Wallace *et al.* 2004b). Terahertz image analysis and processing techniques make it possible to further analyse the interaction between T-rays and biological molecules, cells or tissues, to visualize small particles and opaque objects, and to achieve the recognition of biomedical patterns. These are topics of great importance to biomedical science, biology, and medicines. In this Chapter, three techniques will be reviewed for image analysis that utilize THz radiation: terahertz pulsed spectroscopy, image recovery using THz-TDS techniques, and terahertz identification systems, all of which have a variety of possible applications in the biomedical field.

4.1 THz spectroscopy for biomedical signal identification

Imaging is a main application of THz-TDS. It is valuable for imaging applications not only for observing the structure of the sample but also for performing spectroscopic analysis and identification (Wang *et al.* 2003, Beard *et al.* 2002).

4.1.1 Time-resolved terahertz spectroscopy

Since THz-TDS is a time-domain technique, timing extraction is viewed as a simple but important compression step, for example, by finding the time-delay of the waveforms after passing through materials. This is applicable in assessing thickness variations, in determining the position of unknown objects in reflection geometry, etc. (Mittleman *et al.* 1996).

The non-contact measurement of burn depth and severity using terahertz radiation is a topic of interest. Uncooked chicken breast is commonly used as a test sample for observing the interaction with terahertz radiation in burn diagnostics research. In the experiment represented by Mittleman *et al.* (1996), the reflected THz pulse trains are measured at four different points with various levels of tissue damage on the sample. The results show that a reflected pulse arises from a (buried) interface between tissue layers with and without water content. Since reflected pulses contain a great deal of information about the dielectric properties of the damaged tissue, terahertz radiation can be viewed as a potentially valuable diagnostic technique in biomedical optics.

Pulsed terahertz (THz) wave technology has also been applied for biosensing. An affinity biosensor has been reported using terahertz pulsed radiation in the far-field to monitor binding between biotin and avidin molecules (Menikh *et al.* 2004). Amplified detection of biotin-avidin binding is obtained on supported membranes composed of biotin layers on a quartz surface, which is treated with octadecanol. A conjugation consisting of agarose particles and avidin is applied. Through dithering the quartz slide holder in the THz beam, the resultant THz difference signal between biotin and the biotin-avidin complex is found to increase dramatically. This is due to the increased contrast of refractive index resulting from the chemical binding, which enhances the difference signal (Ferguson and Zhang 2002, Menikh *et al.* 2002). As avidin has a very strong affinity for biotin and is capable of being bound to any biotin-containing molecules, the developed detection technique can potentially be used to detect DNA

hybridization and antigen-antibody interactions (Ferguson and Zhang 2002). A similar technique is applied to DNA near-field imaging (Brucherseifer *et al.* 2000).

Equally important, the pulsed and time-resolved THz-TDS technique is capable of identifing and separating standing waves from the fundamental signal in the spectrometer. This is because some molecular vibrations can be observed via terahertz radiation at room temperature. Fischer *et al.* (2005b) suggest to distinguish two artificial RNA single strands, composed of polyadenylic acid (poly-A) and polycytidylic acid (poly-C) via the different transmission spectra of the two measured terahertz responses. They observe that the poly-C sample shows stronger signal attenuation than the poly-A sample.

4.1.2 Frequency depend terahertz spectroscopy

In the terahertz region, many molecules have a characteristic 'fingerprint' absorption spectra (Mittleman *et al.* 1996). At frequencies below approximately 6 THz, substances in the condensed phase are held together by either ionic, covalent or electrostatic forces, and therefore the lowest frequency modes will be associated with intermolecular motion (Fischer *et al.* 2005a). The interaction between THz radiation and biological molecules, cells, and tissues can be understood as a classical electromagnetic (EM) wave interaction with materials. The medium is characterized in terms of its permittivity ε (the ability of the medium to be polarized) and conductivity σ (the ability of ions to move through the medium). On the other hand, at higher frequencies, transitions between different molecular vibrational and rotational energy levels become increasingly important and are more readily understood using a quantum-mechanical framework (Smye *et al.* 2001). Terahertz pulse spectroscopy provides information on low-frequency intermolecular vibrational modes (Taday *et al.* 2003). In this Section, the vibrational modes of the isolated molecules at low terahertz frequencies are reviewed with related to several different substances and structures, including polarized water, tissue, DNA strands and retinal molecules. Meanwhile, crystalline states of a drug are investigated for pharmaceutical applications.

The absorption and detection in water, liquid and tissue

Terahertz radiation interacts strongly with polar molecules, a prime example being water. Water molecules absorb terahertz waves, on the one hand limiting penetration of the radiation in moist substances, and on the other hand making it readily

detectable even in very low concentrations (Mickan *et al.* 2000). Polar water molecules are active in the infrared region and have various vibrational modes (Pickwell and Wallace 2006). The top of Fig. 4.1(a) shows the vibrational mode of the gas phase of water molecules. In the mid- to far-infrared, the vibrations involve combinations of the symmetric stretch ($\nu 1$), asymmetric stretch ($\nu 3$), and bending ($\nu 2$) of the covalent bonds. The vibrations of water molecules are thought of as restricted rotations, a rocking motion, shown on the bottom of this figure. In liquid water, since hydrogen bonds are much weaker than the covalent bonds (intra-molecular), their bond lengths are much longer (1.97 Å versus 0.96 Å), which is shown in Fig. 4.1(b). This bond state model plays an important role in the understanding of water molecule modes at THz frequencies.

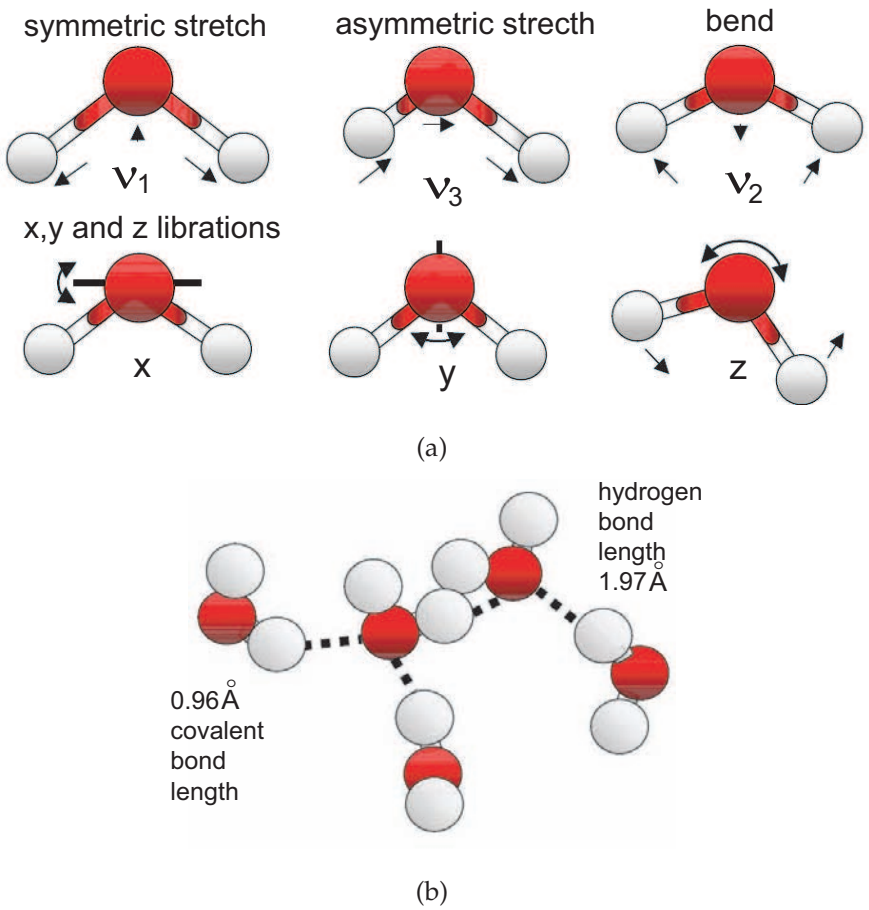


Figure 4.1. Water molecule vibrational moles. (a) The main vibrational modes in water. (b) A schematic diagram illustrating the differences between intra- and inter-molecular bonding in water. After Pickwell and Wallace (2006).

Studying liquids, such as water, with THz time-domain spectroscopy determines the real and imaginary components of the permittivity, which is then related to the intermolecular dynamics. A Debye relaxation model can be used to analyze the strong absorption of terahertz radiation in polar liquids at least up to 1 THz (Rønne *et al.* 1999, Siegel 2004).

Solvation and liquid dynamics play a very important role in essentially all liquid phase processes. Solvent effects are significant in photoexcitation, photoionization, electron transfer, proton transfer, energy dissipation, and liquid phase reactions (*American Physical Society* 2007). Probing the low-frequency solvent modes directly is a proposed method for understanding the solvent response to the above processes. THz time domain spectroscopy has been applied to explore low frequency intermolecular dynamics and solvation in liquids (McElroy and Wynne 1997, Haran *et al.* 1997, Kindt and Schmuttenmaer 1999, Giraud and Wynne 2003).

Biological tissue is generally dominated by polar liquids or preserved in polar liquids. Due to the exceptionally high absorption losses of polar liquids at terahertz frequencies, it is impossible for terahertz radiation to penetrate through biological materials of any substantial thickness. However, the same high absorption coefficient that limits penetration in tissue also promotes extreme contrast between substances with less or higher degrees of water saturation (Siegel 2004). This property has proven advantageous in the examination of the properties of water uptake and distribution in plants (Hadjiloucas *et al.* 1999, Hadjiloucas and Bowen 1999), as well as in the severity of burns on necrotic skin samples (Mittleman *et al.* 1999). In addition, Woodward *et al.* (2003) and Wallace *et al.* (2004a) describe the application of TPI techniques for imaging basal cell carcinoma (BCC) *ex vivo* and *in vivo*. The BCCs show an increase in absorption of terahertz compared to normal tissue. The observed increase in terahertz absorption is most probably due to either an increase in interstitial water within the diseased tissue (Lahtinen *et al.* 1999) or a change in the vibrational modes of water molecules with other functional groups. More research using terahertz radiation in tissue identification has been reviewed by Siegel (2004).

Molecular signatures for the identification of full compounds

As for the identification of full compounds using molecular signatures with terahertz time-domain systems, there exist several difficulties. The difficulties lie in the inherently broad spectral signatures, relatively weak differential absorption compared to

low pressure gasses. Nevertheless, there are a growing number of multiply confirmed observations regarding particular resonant signatures for the identification of full compounds (Choi *et al.* 2002).

It especially becomes interesting in recent years to observe the conformational structure, binding states, and vibrational or torsional modes of proteins and oligonucleotides (Markelz *et al.* 2000, Martel *et al.* 1991) via analyzing spectral features in the terahertz range (Siegel 2002). The reflection or absorption signatures of a sample can afford the information about chemical or physical changes that the sample undergoes. These changes may involve a difference in conformational state, a change of density or polarizability, dehydration, or denaturing or a temperature shift. The comparison of these signatures of sample spectra is the most reliable method to achieve the identification of target objects (Siegel 2004).

Nishizawa *et al.* (2005) illustrate a widely tunable coherent terahertz scanning system to achieve an automatic measurement for terahertz (THz) transmission spectroscopy using the samples consisting of nucleobases, nucleosides, deoxynucleosides, and nucleotides, all of which are the components of RNA and DNA molecules. THz spectra of those samples are measured in crystalline states in the range of 0.4-5.8 THz. The experimental results show that the molecules have quite different characteristic spectral patterns in this frequency region and the patterns are sufficient for identifying and discriminating these molecules.

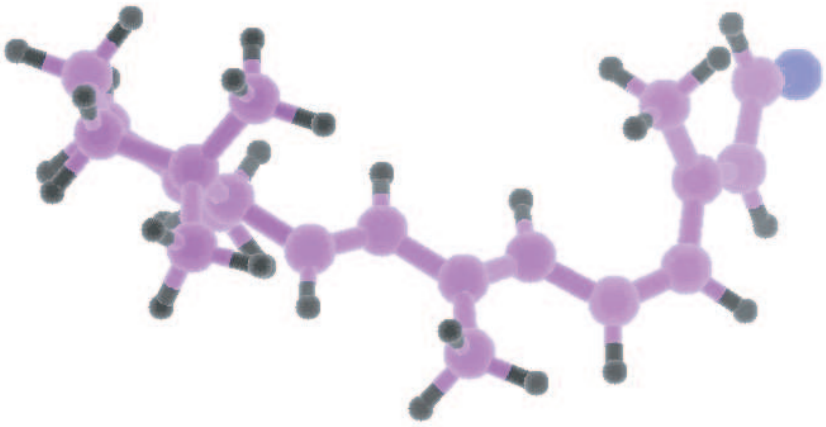
Pulsed terahertz spectroscopy was first reported by Markelz *et al.* (2000) for examining the low frequency collective vibrational modes of biomolecules, i.e. DNA, Bovine Serum Albumin and Collagen between 0.1 and 2.0 THz. It is observed that broadband absorption increases with frequency and a large number of the low frequency collective modes for these systems is suggested to be IR active. Herrmann *et al.* (2005) carry out the measurements of THz spectra of Poly(dA-dT)-Poly(dT-dA) DNA and Poly(dG)-Poly(dC) DNA based on the THz complex refractive index. The resultant spectral features show that those samples are distinguishable in the range from 0.1 to 2.4 THz. Additionally, the absorption spectra of lactose, aspirin, sucrose and tartaric acid are formed in THz pulsed transmission mode by Fischer *et al.* (2005a). The results show the four different chemicals contained in the sample are clearly identifiable.

Long-wave absorption features are intrinsic properties of biological materials, which are determined by phonon modes. The direct identification of terahertz phonon modes in biological materials is very challenging. Globus *et al.* (2002) use Fourier-transform

infrared spectroscopy to investigate the phonon modes in DNA macromolecules via studying the structure of transmission and absorption spectra. Different DNA samples display multiple modes with sequence-specific uniqueness. It provides experimental evidence confirming DNA samples to possess multiple dielectric resonances, via their submillimeter-wave spectra (i.e. $\sim$ 0.01-10 THz), and establishes the initial foundation for the future use of submillimeter-wave spectroscopy in the identification and characterization of DNA macromolecules. Though the above study is based on Fourier-transform infrared spectroscopy, it motivates the future research of phonon modes in DNA macromolecules using terahertz time domain spectroscopy.

More terahertz research groups in Germany and Australia, have worked on the terahertz spectrum regarding the study of the photoisomerization of the retinal chromophore (Walther *et al.* 2000, Jones *et al.* 2006). The conjugated polyene chain of the biologically important chromophore retinal in low-frequency torsional vibration modes is investigated. The absorption and dispersion spectra of three different retinal isomers (all-trans; 13-cis; and 9-cis retinal) in the far-infrared region between 10 and 100 cm^{-1} (0.3 ± 3.0 THz) are measured by terahertz time-domain spectroscopy at 298 and 10 K. At low temperatures it is observed that the broad absorption bands resolve into narrow peaks correlating to torsional modes of the molecule. The modes within the molecule can be approximately localized by the comparison of the absorption spectra of different retinal isomers. Fig. 4.2(a) illustrates a retinal isomer visualized in Molden (Schaftenaar and Noordik 2000), optimized with Gamess (Guest *et al.* 2005). Fig. 4.2(b) shows the mammalian retina with rod cells.

A very promising application for terahertz molecular spectroscopy in the pharmaceutical industry is being pursued by Teraview Ltd., Cambridge, U.K., for detecting the presence of unwanted polymorphs in prepared drugs (Taday *et al.* 2003, Strachan *et al.* 2004, Strachan *et al.* 2005). The crystalline structure of the Polymorphs of Ranitidine Hydrochloride is investigated using terahertz pulse spectroscopy. The considerable differences in the THz spectrum between two different forms of polymorphs are especially observed at around 1.10 THz. It turns out to be the obvious identification of the two structures of Ranitidine Hydrochloride. More work on the identification of crystalline structure drugs using terahertz pulsed spectroscopy has been reviewed by Zeitler *et al.* (2007b).



(a)

NOTE:
This figure is included on page 60 of the print copy of the thesis held in the University of Adelaide Library.

(b)

Figure 4.2. Retinal isomer. (a) Illustration of a retinal isomer visualized in Molden (Schaftenaar and Noordik 2000), optimized with Gamess (Guest *et al.* 2005). (b) Illustration of the mammalian retina with rod cells. The stacks of membrane discs (middle) consists of the outer segment of a cell with protein rhodopsin involved. A retinal chromophore in its 11-cis conformation is embedded in the rhodopsin molecule (right). The vision process is eventually triggered by the isomerization of the retinal to the all-trans configuration. After Walther (2003).

NOTE:
This figure is included on page 61 of the print copy of the thesis held in the University of Adelaide Library.

Figure 4.3. Time-frequency domain features of THz signals. Contour plots of the real part (left) and imaginary part (right) after conducting the db4 continuous wavelet transforms of terahertz measurements corresponding to: (a) leather, (b) lycra and (c) the difference. After Galvão *et al.* (2003).

4.1.3 Time-frequency domain features of terahertz signals

Time-frequency methods have been proven to hold great promise for optimizing the extraction of the spectroscopic information contained in each terahertz pulse, for the analysis of more complex signals using THz-TDS (Berry *et al.* 2005). The wavelet transform (WT) is a popular technique suited to the analysis of short-duration signals, especially signals with sudden and unpredictable changes that often carry the most interesting information (Qian 2002). The WT decomposes time domain signals into components of different magnitudes and position, which makes possible to modify wavelet transform coefficients and keep the coefficients reflecting the relevant information (Meyer-Base 2003). The robustness of wavelet transform (WT) can be applied for the compressed measurements (Galvão *et al.* 2003) in terahertz pulsed imaging.

Currently, there are a few terahertz groups in the world which are performing the research using wavelet transforms to achieve the analysis of terahertz pulsed responses. In the University of Leeds, a wax embedded histopathological melanoma section and a tooth slice are measured for the discrimination of cancer tissues from normal cells, and classification between dentine and enamel, respectively. The separations are achieved with acceptable error (Berry *et al.* 2005). At the University of Reading, leather and lycra are used as samples to study data compression in terahertz pulsed imaging (TPI). It is found that the estimates of refractive index and absorption coefficients are not

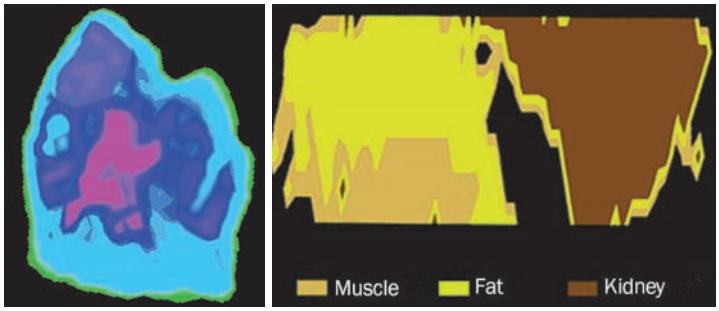
significantly altered when the terahertz data are reconstructed from only 20% of DWT coefficients (Galvão *et al.* 2003). Fig. 4.3 shows the contour plots of the real part (left) and imaginary part (right) after conducting the db4 continuous wavelet transforms of terahertz measurements corresponding to: (a) leather, (b) lycra, and (c) the difference. The reconstructed signal after wavelet transforms shows improved classification performance. Further examples exploiting the time-frequency domain for classification of signals from biomedical samples will be reviewed later in Section 4.3.

4.2 2D and 3D terahertz biomedical imaging

The Rayleigh scattering of light is a major difficulty hindering image processing in the terahertz regime (Han *et al.* 2000) and other optical based techniques. In order to improve the image quality and overcome the Rayleigh scattering at greater depths, it is expected to increase the wavelength of light used, since the degree of Rayleigh scattering is inversely proportional to the fourth power of the wavelength, λ (Ferguson and Abbott 2001a). However, relatively long wavelengths limit the spatial resolution of the objects. In order to obtain higher resolution, a bright source of radiation at intermediate wavelengths is needed (Reid and Fedosejevs 2005). In other words, the wavelength has to be sufficiently small to provide good resolution, yet large enough to prevent serious losses by Rayleigh scattering. Though Rayleigh scattering still remains an obstacle in many biomedical applications, recent advances have sought to address these effects (Pearce and Mittleman 2003). It is expected that terahertz radiation should produce less scattering than visible and near-infrared frequency, which is a major advantage for terahertz imaging (Berry *et al.* 2005). It has been borne out by experiments that terahertz reveals generally higher image contrast than near-infrared, although the near-infrared pulses posses high power (Cogdill *et al.* 2007).

As mentioned in Chapter 2, terahertz pulsed imaging can be achieved via a pump-probe detection system. In pump-probe detection, terahertz pulses are focused on a particular point on an object. The transmitted or reflected portion of each pulse is then detected after a time delay. By measuring the length of this delay, a 3-D image of the internal structure can be discerned. The target image can be built up via scanning the beam, or reflecting from or transmitting through the sample at each position of the sample. A variety of *in situ* measurements in various applications can be carried out. Alternatively, a terahertz snapshot of the object can be achieved via using a

charge-coupled device (CCD) for the measurement involving a moving target or source (Arnone *et al.* 2000).



(a) (b)

Figure 4.4. Pulsed THz imaging of a human tooth and a sample of pork. (a) An image of a human tooth formed from terahertz radiation in transmission mode. (b) Illustration of the terahertz pulsed image of a sample of pork. After Arnone *et al.* (2000).

A major advantage of terahertz imaging is in its diagnostic capabilities. Fig. 4.4(a) is an image of a human tooth formed from terahertz radiation in reflection mode. Due to the different mineral content among the different layers of the tooth, a large difference in the index of refraction can be observed for the different types of tissues. The distance between the different layers can be resolved readily from the reflected terahertz radiation by measuring the time of flight. Furthermore, the spectral information from terahertz pulses are used to distinguish different types of soft tissues, such as muscle, fat, and kidney, because each has a characteristic 'fingerprint'. Fig. 4.4(b) illustrates the terahertz pulsed image of a sample of pork. The absorption spectrum at each pixel highlights the regions of muscle and fat in detail.

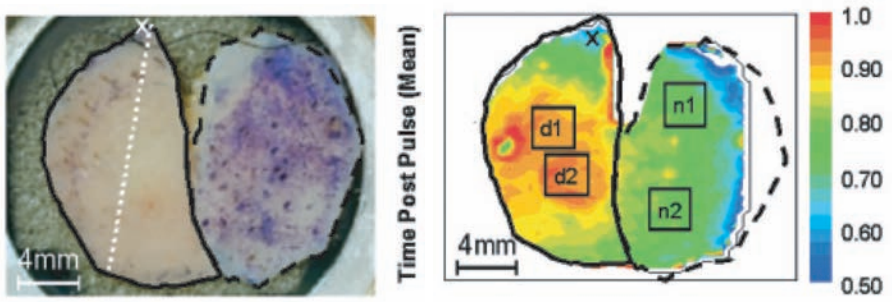
4.2.1 Cancer cell detection

Since water has strong absorption at THz frequencies and tumours tend to have different water content from normal tissue, a terahertz image can be achieved by using a contrast mechanism based on variations in water content.

Skin cancer

Terahertz pulse imaging of *ex vivo* basal cell carcinoma was reported first by the research group in the University of Cambridge (Woodward *et al.* 2002). Basal cell carcinoma is the most common form of skin cancer. The terahertz pulsed image is generated

in reflection mode using a terahertz time-domain analysis technique (Woodward *et al.* 2003). The recovered terahertz image is shown in Fig. 4.5. The diseased tissue shows an increase in absorption compared to normal tissue, which is attributed to either an increase in the interstitial water within the diseased tissue or a change in the vibrational modes of water molecules (Woodward *et al.* 2002, Woodward *et al.* 2003). This result from Cambridge University motivates great interests in skin cancer detection (Wallace *et al.* 2004a, Woodward *et al.* 2004, Pickwell *et al.* 2004) and makes a valuable exploration using terahertz radiation.



(a) (b)

Figure 4.5. Skin cancer THz detection. A comparison is conducted between the visible image (a) and the THz image (b). The diseased tissue, on the left of (a), is marked by a solid boundary, the normal tissue on the right by a dashed boundary. The dotted white line indicates the axis of the vertical histology section. The white 'x' marks the location of the suture. In (b), d1 and d2 label the diseased tissue, with equal size areas, and n1 and n2 label the normal tissue. After Woodward *et al.* (2003).

Breast cancer

The feasibility of using terahertz pulsed imaging to map margins of exposed breast tumors is investigated by Fitzgerald *et al.* (2006). The size and shape of tumour regions on terahertz images are compared with those identified at histopathologic examination of the imaged section. The minimum of the terahertz impulse function ($E_{\min}$) as parameter is used to map the target measurement for the two dimensional image. The THz pulsed image is achieved in reflection mode, as illustrated in Fig. 4.6. It is observed that photomicrographs contained many more pixels than the terahertz images. Though, this resultant experiment demonstrates the potential of terahertz pulsed imaging to depict invasive breast carcinoma *in situ* under controlled conditions.

Figure 4.6. Breast cancer THz inspection. The shape of the tumour regions is compared via photomicrographic scanning (a) and terahertz scanning (b) using a invasive ductal carcinoma as a sample. The tumour regions in (a) are bounded by the black outline. In (b), the terahertz image is mapped via parameter $E_{\min}$. A threshold is applied to delineate regions (red) for (c) photomicrograph, and (d) terahertz image. Numbers of pixels in the cancer regions have been summed in x and y directions to form profiles in (e) and (f), respectively. The profiles consisting of circles are from a photomicrograph, and black dots from a terahertz image. After Fitzgerald *et al.* (2006).

4.2.2 Brain section detection

Terahertz quantum cascade lasers as bright laser sources have been shown to be promising to image biomedical samples. It was first demonstrated by Darmo *et al.* (2004) to image a specimen of dehydrated rat brain by using a terahertz QCL. This new optoelectronic source of coherent terahertz radiation allows researchers to build a compact imaging system with a large dynamic range and high spatial resolution. The image of a brain section is obtained at 3.4 THz, and illustrated in Fig. 4.7. Distinct regions of brain tissue rich in fat, proteins, and fluid-filled cavities are resolved showing the high contrast of terahertz radiation for biological tissue. The Fabry-Pérot modulation of the background in the image is suppressed by tilting the sample holder. The THz images

NOTE:
This figure is included on page 66 of the print copy of the thesis held in the University of Adelaide Library.

Figure 4.7. Rat brain image using a THz QCL. White-light (a,b) and THz (c,d) images of rat brain frontal sections. The samples are about 30 μm thick and fixed onto a gold-coated flat mirror. The pixel size is 200 μm $\times$ 200 μm. The dark (high absorption) structures in THz image are from white matter tissue (e.g., corpus callosum, hippocampus, capsula interna, commisura anterior) which are more or less heavily myelinated and consisted largely of lipids, i.e. fats. The different fat content thereby determines the differing grey values. The grey matter of the brain (e.g. cortical cortex), which naturally has a much higher content of water and proteins, gives higher signal intensities in THz images (less absorption) than do tissues with high fat content. The brightest structures (the lowest THz absorption) in the images correspond to the ventricular system, which *in vivo* contain water and *in vitro* contain air. After Darmo *et al.* (2004).

are collected in transmission mode at scanning step size of 200 μm. The continuous-wave terahertz imaging can provide valuable data on biological samples via carefully selecting possible T-ray wavelength. These data appear complementary to those obtained from white-light images.

4.2.3 Tablet coating detection

Three dimensional terahertz pulsed imaging is evaluated as a novel tool for the non-destructive characterization of different solid oral dosage forms (Zeitler *et al.* 2007a).

The measurements using 3D TPI techniques are fast and fully automated and allow to determine the spatial and statistical distribution of coating thickness in single and multiple coated products. In soft gelatin capsules, TPI is able to measure the thickness of the gelatin layer and to characterize the seal between the gelatin layers for quality control. The resultant mapping of a gelatin capsule from practical measurements based on reflective mode TPI is shown in Fig. 4.8, which is obtained from experiments performed by Zeitler *et al.* (2007a). The 3D THz pulsed imaging technique has the potential for much wider applications in the process analysis field.

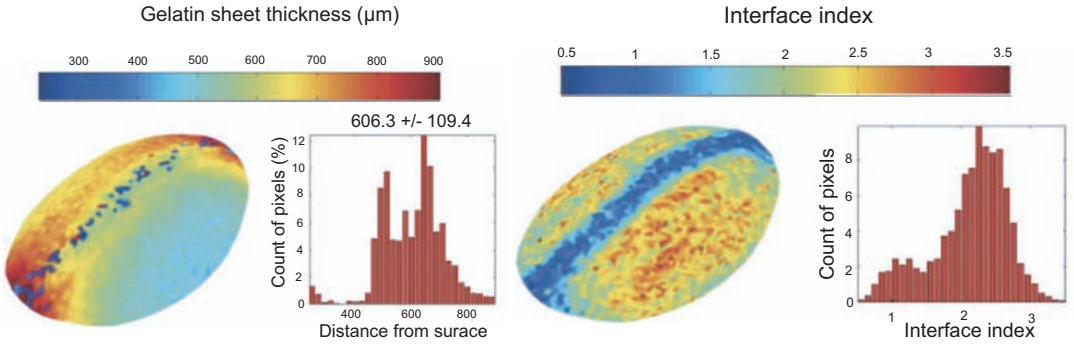


Figure 4.8. Thickness measurements of soft gelatin capsules using THz 3D imaging. Spatial distribution and histogram of the gelatine layer thickness (a) and interface index (b) over the surface of the capsule. The time delay between the two peaks allows the thickness calculation of the gelatin sheet for every single recorded pixel, shown in (a). One of the two peaks is the signal from air-gelatin interface, and the other is from gelatin-liquid interface. It reveals that the thicknesses of the two soft gelatin sheets are different, which match to the original thickness of the target capsule. By illustrating the relative peak strength of the reflection signals in the terahertz waveform rather than the time delay, the area of the seam between the two gelatin sheets is clearly distinguished from the rest of the capsule, shown in (b), which is much thinner than the two gelatin sheets themselves. After Zeitler *et al.* (2007a).

4.3 Pattern recognition of biomedical samples

Automatic classification of data is a significant application of imaging techniques in the signal processing research area. Such systems need to be capable of analyzing and dealing with measure image statistics and automatically producing a correct inspection outcome (Meyer-Base 2003, Yin *et al.* 2007b).

THz images are parametric images derived from the measured pulses (Löffler *et al.* 2002, Yin *et al.* 2007g). For instance, pulse height and shape, and delays in the time domain carry contrast information. Additionally, the spectral content of the pulse can be exploited for pixel by pixel spectroscopic analysis and material characterization, including transmittance and absorbance parameters, as commonly used in Fourier transform infrared spectroscopy (Lasch and Naumann 1998, Withayachumnankul *et al.* 2005). Combinations of parameters are useful for developing thickness-independent parameters that allow differentiation between time delays due to refractive index variations or sample thickness. In many cases, the parameters can be compared to a reference pulse to obtain a relative measure (Fitzgerald *et al.* 2002). These parameters are significant in pattern recognition for feature extraction. Via feature extraction, several critical features of material from the complex T-ray spectral responses can be separated to acquire a good classification performance (Yin *et al.* 2007h, Bow 2002).

To extract features and form accurate classification, it is possible to use system identification to estimate a system that best describes measured data. In order to estimate a system, many methods suggest estimating linear model coefficients (Kalouptsidis and Theodoridis 1993). These methods applied to terahertz imaging research mainly include (i) *k*-means classification of skin cancer cells by applying multispectral clustering techniques (Berryman and Rainsford 2004); (ii) the separation of the component spatial patterns of chemical samples (Watanabe *et al.* 2003); (iii) the Euclidean distance classifier by continuous wavelet transforms of THz transient spectrometers (Galvão *et al.* 2003); (iv) support vector machines (Withayachumnankul *et al.* 2006) for the classification of bone cancer cells. The following sections review the basic classification principles and pattern processing techniques for terahertz pattern recognition. The suggested parameter extraction methods along with linear and nonlinear classifiers mentioned above are discussed.

4.3.1 Extracted parameters for terahertz pattern recognition

Parameter terahertz images for pattern recognition are well illustrated by Löffler *et al.* (2002). This terahertz group map 2D images by light- and dark-field TPI techniques with a thin-cut canary's head as a sample. Both time- and frequency-domain methods are applied for terahertz parameter extraction. This dark-field TPI technique is to explore the image mapping by using the deflected terahertz beams beyond the normal beam-propagation direction, due to either diffraction or scattering effects in the sample

(Löffler *et al.* 2001). Different parameter extraction methods in the time- and frequency-domain are capable of achieving successful tissue classification after applying a suitable threshold (Ferguson *et al.* 2002c, Löffler *et al.* 2002).

4.3.2 Multispectral classification for terahertz pulsed imaging

The concept behind multispectral classification techniques is borrowed from multispectral satellite imaging (Vohland *et al.* 2007). It allows labelling multispectral data in terms of corresponding different physical properties of material and display of a coloured multispectral image. Different colours overlaid in multispectral images represent different classes in which the image is classified (*Defence Research & Development Organization* 2008). Multispectral classification composes of supervised and unsupervised learning. The techniques are efficient for classifying heterogeneous objects (Zhang *et al.* 2004) and is well-suited to the analysis of terahertz pulsed imaging data (Ferguson *et al.* 2002c, Berryman and Rainsford 2004). Multispectral classification algorithms are commonly realised by applying current image processing software systems (*University of California* 2008).

ISODATA and k-means classifications

ISODATA and k-means classifications are two most frequently used methods for unsupervised classification, which means training data are not involved for the classification procedure (Bow 2002). The two techniques were first employed in THz imaging by Berry *et al.* (2004). Both of these algorithms are iterative procedures. The aim is to repeatedly classify each pixel to the closest cluster till the change between the cluster mean vectors is small. The iterative procedures start from an arbitrary assigned initial cluster vector. The 'change' can be defined either by measuring the distances the mean cluster vectors have changed from one iteration to another or by the percentage of pixels that have changed between iterations (*Yale University* 2008).

The ISODATA algorithm is similar to the k-means algorithm with the distinct difference that the ISODATA algorithm allows for different number of clusters and some further refinements by splitting and merging of clusters (Duda and Hart 1973). In contrast, the k-means assumes that the number of clusters is known *a priori* (MacQueen 1967), and then attempts to minimise the total intra-cluster variance.

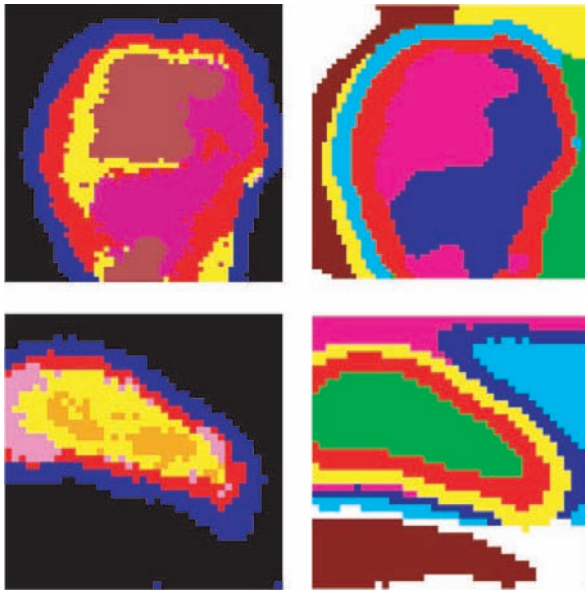


Figure 4.9. Multispectral clustering for THz classification. The resultant separated patterns are based on two specimens: BCC on the top; melanoma on the bottom. Boundary regions, are shown in a shade of blue. The sub-classifications are achieved within the innermost regions for both BCC and melanoma. The ISODATA based images are shown on the left hand. Results of K-means clustering into eight classes using the 64 time series values are shown on the right hand. After Berry *et al.* (2004).

Fig. 4.9 shows the visualized multispectral clustering techniques based on the experiment conducted by Berry *et al.* (2004). Pulsed terahertz imaging in transmission mode is used for terahertz measurements. Two specimens are examined for study, one is basal cell carcinoma (BCC) (on the top) and the other is melanoma (on the left bottom). Cells are treated by dehydration and fixed with formalin, for the sample stability. The unsupervised ISODATA algorithm is used for classifications resulting from three parametric images with two merge steps. The classified images from the ISODATA algorithm are shown on the left hand side of Fig. 4.9. The internal structure of the BCC is more organized than that of the melanoma, which is consistent with the appearances of the original target slice. Results from a *k*-means classification using time series features appear on the right hand side of Fig. 4.9. These show the same shapes as the ISODATA classification using three features. For the BCC, the well-defined internal structures are also shown. There are more classes defined in the background regions of these images because the number of classes is fixed at eight, without the merging options used for the ISODATA algorithm.

Principal component analysis

Principal components analysis (PCA) is a technique used to reduce multidimensional data sets to lower dimensions for analysis (Jolliffe 2002). Principal component analysis algorithms are employed by Shikata *et al.* (2002) and Watanabe *et al.* (2003) to demonstrate the separation of the component patterns of chemical samples in multispectral THz images. The experiment setup uses a widely tunable coherent THz-wave parametric oscillator (TPO) (Kawase *et al.* 1996). The spatial patterns consists of two components: Palatinose and 5-aspirin, both of which are fixed on a sample holder. The holder is mounted on a x-y linear moving stage. All the samples are pressed into pellets with thickness of 1.0 mm. The content is analyzed via applying five transillumination THz spectroscopic images: palatinose and 5-aspirin are prepared in three (50%, 40%, and 20%) and two (50% and 20%) concentrations, respectively. The frequency range covers from 1.3 to 1.8 THz in 0.1 THz steps.

The intensity of terahertz images keeps attenuated by absorption linearly, which is caused by the concentration and the thickness of the chemical samples. Therefore, the image $F(x, y, f)$ observed at each of the f THz is expressed as a linear combination of absorbance spectra $S(f)$ and spatial pattern $P(x, y)$, which satisfies a matrix equation $[F] = [S][P]$ for principal component analysis. Since $[F]$ and $[S]$ are measured directly, a least-squares method allows the matrix equation to be solved. As a result, the matrix $[P]$ is described by $[P] = ([S]^{\mathrm{T}}[S])^{-1}[S]^{\mathrm{T}}[F]$, where T labels the matrix transpose operation.

Fig. 4.10(a) shows the measurement of the spectral data of well separated samples of 50% 5-aspirin and palatinose using the TPO, which corresponds to the one of the columns of matrix $[S]$. A THz image at 1.5 THz is displayed in Fig. 4.10(b), which corresponds to one of row vectors of matrix $[F]$. The resultant component analysis is shown in Fig. 4.10(c) and (d). The estimated values of concentrations with respect to palatinose and 5-aspirin approximately agree with the known concentrations.

Though the principle component analysis and linear discriminant analysis (LDA) are commonly used for feature extraction (Jolliffe 2002, McLachlan 1992), these methods do not offer a universal panacea. It has been recognised that there exists the problem of 'overfitting', where complex classifiers such as neural networks result in overly detailed decision boundaries. The boundaries described match the training data too closely and thus fail to generalise a accurate classification. It is the subject of extensive research (Domingos 1999, Cubanski and Cyganski 1995, Fleming 2001).

NOTE:
This figure is included on page 72 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 4.10. Principal component analysis of THz spectral. (a) Spectral data of 50% 5-aspirin and palatinose. (b) THz images at 1.5 THz. Five mixtures are involved: 50%, 40%, and 20% of 5-aspirin mixed with palatinose plus 50% and 20% of palatinose mixed with 5-aspirin. The scale of the image is the logarithm of the observed image intensity. Component spatial patterns with different calculated concentrations of (c) palatinose and (d) 5-aspirin. After Watanabe *et al.* (2003).

4.3.3 Classification of THz spectra in the wavelet domain

The performance of a classifier based on the output of a filter bank was first illustrated by Galvão *et al.* (2003) in terahertz optical research. It turns out to be considerably better than that of the Euclidean distance classifier in the original spectral domain. An optimal feature extraction in the wavelet domain is performed, by maximizing an objective function. The object function is achieved by dividing the power of the wavelet coefficients by the noise amplification factor. The procedure is capable of enhancing

NOTE:
This figure is included on page 73 of the print copy of the thesis held in the University of Adelaide Library.

Figure 4.11. Optimal THz classification in the wavelet domain. Graphical illustration of the transformations regarding the optimization process, by representing the discrimination of two responses at four particular time, or two points in two four-dimensional spaces. Considering the left half of the vectors after performing the Fourier transformation of responses, each of the spectra pattern consists of real and imaginary parts at two frequencies ω_1 and ω_2. In (a) and (b), the two patterns indicated by two points are discriminated in two complex planes, each plane associated with one of the frequencies. The noise associated with the patterns is represented by the larger circles. The noise is equal in all directions because the Fourier transform is orthogonal. The difference ΔX_a as a vector is represented in each plane ((c) and (d)). The differences in complex insertion loss ΔL are understood as ratioing against the background at each frequency, which is equivalent to rotating and contracting the difference vectors ΔX_a. Since the background intensity is different at each frequency, the noise level becomes different in $\Delta L(\omega_1)$ and $\Delta L(\omega_2)$. The result is that the noise is no-longer equal in all directions, as shown in (e) and (f), which turns out to be a nonoptimal Euclidean distance classifier (Raudys 2001, Hadjiloucas *et al.* 2002). The optimization of the objective function is to find a single direction for the k^{th} wavelet coefficient by performing a further rotation of the four axes, with an aim to achieve the maximal projection OB of the difference with respect to the projection AC of the noise. After Galvão *et al.* (2003).

the contrast or classifying spectra acquired by either continuous wave or THz transient spectrometers for increasing the dynamic range of THz imaging systems. This optimal feature extraction is graphically demonstrated in Fig. 4.11.

4.3.4 Support vector machines for classification of the terahertz relevant frequencies

In addition to the classifiers mentioned above, the nonlinear classifier known as a support vector machine has also been explored by Withayachumnankul *et al.* (2006), for the classification of the T-ray response of normal human bone cells and human osteosarcoma cells. The magnitude and phase responses within a reliable spectral range are the extracted feature vectors, and a trained support vector machine allows correct classification of the two cell types with a consistent classification accuracy of 89.6%, with only one fifth of the original features retained in the data set. Feature subset selection methods are used to select only an optimal number of relevant features for inputs. The resultant improvement in generalization performance is achieved, and the selected frequencies are used for further describing different mechanisms of the cells, responding to terahertz radiation. Further investigation using support vector machines will be performed and demonstrated in Chapter 9.

4.4 Chapter summary

THz imaging analysis plays an important role in THz pattern recognition. The intention of the review represented in this Chapter is to provide a motivational setting for further exploration of THz material identification, as carried out in this Thesis. These experimental results reviewed in this Chapter are from work performed by international terahertz groups, aimed at highlighting the potential of the application of various signal processing techniques and classification algorithms on THz spectroscopy and pattern recognition. In the next Chapter, we discuss a pattern recognition framework for application to terahertz radiation. We also point out the importance of signal processing techniques for the THz identification system, especially the exploration of feature extraction methods will result in a successful differentiation of THz pulsed responses, which is of essential significance to achieve biomedical related THz classification algorithms.

Chapter 5

Pattern Formation and Recognition Using T-rays

PATTERN formation using terahertz radiation can be defined as the generation of a quantitative or structural description of an object, especially an optically opaque object, via T-rays. It follows that a pattern class can be defined as a set of patterns that share some properties in common. As common pattern properties belong to the same class, it enables us to build different models for discrimination. Pattern recognition is the process of categorising any sample of measured or observed data as a member of a candidate class, several of which may be allowed in each particular problem. For pattern recognition, applications tend to be specific and thus require specialised techniques. In this Thesis, sample responses from multiple terahertz experiments are used for pattern recognition case studies.

5.1 Significance of a terahertz pattern recognition system

Pattern recognition techniques can be applied in many different areas to solve existing problems. The various requirements arising from the process of resolving practical problems motivates and expedites the development of pattern recognition. This is a major reason why this discipline has grown so fast (Bow 2002).

One of important parts of this Thesis is the exploration of different approaches to deal with and identify T-ray patterns. Since the output of the T-ray instruments is in the form of raw numerical data, the pattern can be constructed from images in digital format, records of time series, or measurement vectors describing the state of a physical system.

The function of a THz pattern recognition system is to emulate the recognition ability of human beings, while using T-ray imaging as the primary sensing mode. The procedure involves automatic processing of acquired data and making the decisions of pattern discrimination, effectively and efficiently.

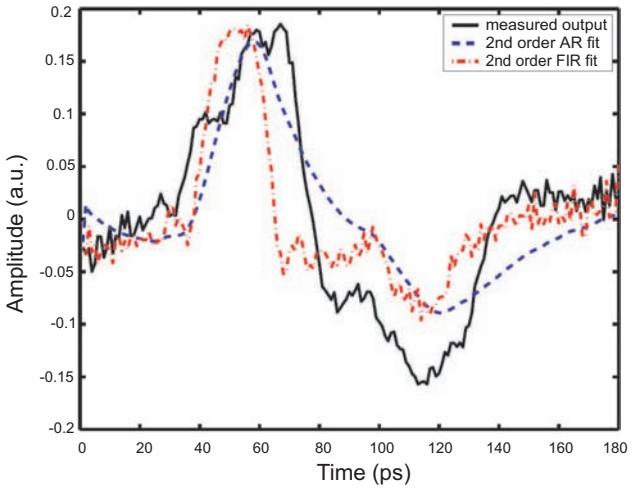
Pattern recognition using T-rays differs from T-ray image processing. T-ray imaging achieves the description of the complete target of interest (Duda *et al.* 2001) via applying entire T-ray imaging measurements. In contrast, T-ray pattern recognition aims to select a fewer number of significant features than in a complete image data set, resulting in simpler representations without incurring a dramatic loss of information. Through pattern analysis, any relations, regularities, or structures inherent in the source of a T-ray response can be found. By detecting significant patterns in the available T-ray relevant data, a T-ray system can expect to make predictions about new THz data coming from the same source (Shawe-Taylor and Cristianini 2004).

Automation of the laboratory examination of routine biomedical signals and images is an important application area for T-ray pattern recognition. Pattern recognition can involve optical property recognition, powder recognition of chemical substances, layer thickness detection of drugs, and DNA sequence identification for example. It may also be possible to screen out abnormal cells, from normal ones, for tissue analysis applications.

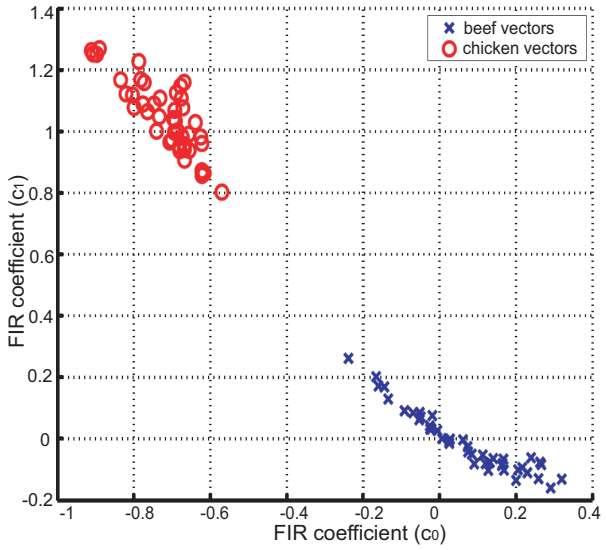
5.2 Mode of the THz pattern recognition system

The pattern recognition system in this Thesis is developed to transform target object information, in the form of T-ray time domain spectroscopic measurements, and then to extract features and recognise patterns corresponding to the underlying sample. This recognition system more easily allows humans to identify the structures or components of the target substances. The difficulties with using THz for this type of work lie in extracting useful information from empirical data. Changes in the THz pulse that are caused by the sample can be difficult to distinguish from those caused by long-term fluctuations in the driving laser source. In order to differentiate between terahertz patterns in data, it is proposed to increase contrast in THz signals and images. For instance, by applying signal processing techniques, the difference in the water content of tissues, in the refractive indices and absorption coefficients of various media can be detected. Alternatively, by tracking suitable molecular resonances, the different optical characteristics possessed by different DNA components can be analysed.

Signal processing techniques have been used to improve the speed, resolution and noise robustness of T-ray imaging systems (Ferguson *et al.* 2002c). These effective processing methods facilitate T-ray signal or image recognition. This Thesis considers a number of signal processing techniques suitable for extracting information from the data obtained in a terahertz pulse measurement system. A sequence of operations are performed on the T-ray responses of target objects in order to achieve desired processing results. For example, in the case of a discrimination of two different substances, chicken and beef, signal processing changes the original THz signal transmitting through the chicken breast, via applying Finite Impulse Response (FIR) and Auto Regressive (AR) fitting, into a modified version of the THz response, shown in Fig. 5.1(a). This modification makes it possible to achieve sample recognition from different classes of chicken and beef sample signals, shown in Fig. 5.1(b).



(a)



(b)

Figure 5.1. Pattern recognition of chicken and beef. (a) Model output for second order FIR and AR filters. The sample holder THz response is used as the model input and the desired output of this model is viewed as the measured response after transmission through the chicken breast. The Wiener-Hopf equation is adopted to calculate the least squares model coefficients for second order Finite Impulse Response (FIR) and Auto Regressive (AR) systems. The model outputs are compared with the desired output. The models are modestly accurate, accounting for 43% and 32% of the actual response respectively. (b) Scatter plot of T-ray transmission for chick and beef. This plot shows the ability of the 2nd order FIR model coefficients to discriminate the two different classes of samples. The optimal FIR model coefficients are calculated on 100 random T-ray responses and plotted. The two classes show a significant difference in their coefficients. After Ferguson *et al.* (2002c).

NOTE:
This figure is included on page 79 of the print copy of the thesis held in the University of Adelaide Library.

Figure 5.2. THz scanner and data recognition system. The block chart shows the three steps of data analysis for pattern recognition. In the data acquisition phase, analog data are collected from a THz receiver, after passing through or reflecting from a target object. The analog data are then converted to digital format suitable for computer processing. Identification of the THz pattern is then achieved at the classification stage. After Federici *et al.* (2007)

5.3 Configuration of the THz pattern recognition system

There are three steps that comprise a THz pattern recognition system: (i) data acquisition through THz imaging, (ii) preprocessing, and (iii) decision classification. Fig. 5.2 illustrates a data recognition system with the use of a THz scanner.

5.3.1 Data acquisition using T-rays

In the data acquisition phase, analog data are collected from a THz receiver, after passing through or reflecting from a target object. The analog data are then converted to digital format suitable for computer processing. THz pulsed imaging setups for data acquisition have been described in Chapter 2.

5.3.2 Preprocessing

During preprocessing, input data vectors are grouped into a set of characteristic features as output. One function of preprocessing is to convert a set of discrete data into a mathematical pattern (feature vectors) so that those data are more suitable for computer analysis (Bow 2002). Data preprocessing and feature extraction are two important steps for carrying out the preprocessing procedure. Data preprocessing isolates the real T-ray responses from the effects of the background noise, and supplies cleaner signals for effective feature extraction.

System and random error are major sources of noise in a terahertz system. One method to reduce the error is to average several measured sequences, but this comes at a cost of time to perform the measurements. Signal processing potentially supplies an improved solution to noise problems. Wavelets are believed to be of critical interest in this research. They possess a range of extremely attractive properties for the denoising preprocessing (Ferguson and Abbott 2001b). Another significant source of error and ambiguity in THz systems is the system hardware itself, including electrical and optical reflections from system components and numerous other effects (Mittleman *et al.* 1996). The process of deconvolution is generally performed to isolate sample signal property from system environment. Since deconvolution is noise sensitive, it can result in a large amount of error when a significant amount of noise is present. Improved methods can optimally remove the noise signal (Ferguson and Abbott 2001b). More system noise from the THz emitter dominates all other noise contributions, which results from random intensity fluctuations of the ultra-fast laser (Haus and Mecozzi 1993, Poppe *et al.* 1998). Other noise sources can be traced back to Johnson and shot noise in the THz detector, as well as thermal background radiation in the THz regime (Duvillaret *et al.* 2000). The noise is incoherent and adds randomly for successive optical pulses, while the signal is coherent and scales linearly with the number of gating pulses (van Exter and Grischkowsky 1990).

Each T-ray measurement with n data points can be viewed as a vector in a n dimensional space, known as a *pattern space*. The time sequence then appears as a point in the pattern space. The object of feature extraction is to reduce dimensionality. It converts the preprocessed data to feature vectors.

By feature extraction, the inherent characteristics found within the measured THz response are acquired, which are used to describe the object, or attributes of the object. THz pulsed imaging is the measurement of a set of time series at each position of a

target object. In this Thesis, the feature extraction operates on a one dimensional signal and produces a shorter feature vector. For THz imaging applications, this analysis is carried out pixel by pixel, on the THz image. These feature vectors are grouped together via a decision function and then are checked to see whether they provide meaningful information. The features of interest can be a series of basic parameters, reflecting the physical characteristics of target materials. They can be derived from the measured THz pulses. Pulse height, shape, delays in the time domain, as well as the spectral content of the pulse in the frequency domain, including transmittance and absorbance parameters, can all provide contrast information. Recognition between the different time delays is available via the combinations of parameters due to refractive index effects or sample thickness.

In this Thesis, the fast wavelet transform is adopted in some instances to achieve effective feature extraction. This wavelet based method especially deals with the assessment of THz process parameters in a given application using the features extracted from the wavelet coefficients of THz measurements (Pittner and Kamarthi 1999). Since the parameter assessment using all wavelet coefficients will often turn out to be tedious or leads to inaccurate results, a preprocessing routine that computes robust features that are representative in correlation with parameters of interest is highly desirable. In this Thesis, the use of Auto Regressive (AR) and Auto Regressive Moving Average (ARMA) models on the wavelet transforms of measured T-ray pulse data are investigated. The features of a processed THz signal are eventually classified by an Mahalanobis distance classifier. The effectiveness of this method is demonstrated via cancer cell discrimination from normal tissue and on the problem of recognising different kinds of powders.

5.3.3 Representation of patterns for machine recognition

The third phase is actually a classifier, which makes a final decision about the cluster of the patterns that have the common features and the discrimination of the patterns that have different properties. The classifier may be linear, piecewise linear, or nonlinear. The data to be classified is the set of feature vectors input from the second phase of preprocessing. The functions used for classification in this Thesis are adjusted during a training phase, guided by knowledge about each pattern. In other words, prior knowledge about correct classification of some data vectors is needed in the training phase of the decision processing.

The classification problem is simply to find a separating surface that partitions the known vectors in to correct classes (Bow 2002). It is expected that the separating surface is able to classify the other unknown patterns by applying the same criterion in the classifier. Since patterns belonging to different classes will cluster into different regions in the pattern space, the similarity between patterns in the feature space can be measured by an inter-pattern distance matrix.

This work carried out in this Thesis is focused on developing feature extraction methods with application to several different classifiers. Since classification is problem dependent, this is carried out in the hope that these pattern recognition algorithms applied in this Thesis, in addition to proving high classification accuracy, may add to our relatively limited knowledge and understanding of the interaction between THz radiation and common biomedical materials.

5.4 Chapter summary

This Chapter discusses a pattern recognition framework for application to terahertz radiation. The design of an identification system emulates the recognition ability of humans, where terahertz radiation is used to obtain the optical parameters of an object. The open question of the identification system is to explore signal processing approaches for automatic processing of terahertz measurements, obtaining effective terahertz pattern recognition. The major difficulties lie in extracting useful information from measured THz data, affected by the long-term fluctuations in the driving laser source. Three strategies are used to realise such an identification system: data acquisition through THz imaging, preprocessing, and decision/classification, where the preprocessing procedure consists of two important steps: data preprocessing and feature extraction. The exploration of feature extraction methods is a focus of this Thesis, with application to several different classifiers and will appear in Chapters 6 to 9. They are successful in the differentiation of THz pulsed responses via applying dimensionality reduction to feature sets. In next Chapter, we present the mathematical basis of wavelet transforms. Wavelet transforms are an important signal processing technique, and the background theory underpins most of the work in this Thesis. It is introduced by the link to Mallat's pyramid algorithm and from the multiresolution point of view. It is also an essential step towards the future goal of developing curvelet based THz classification algorithms, which remains an important open question for future research.

Chapter 6

Wavelet Transforms

THE history of the wavelet transform can be traced back to Fourier theory. The Fourier decomposition expands a signal as an integral of sinusoidal oscillations over a range of frequencies (Ng 2003). A major limitation of Fourier theory is mixing of temporal information in the Fourier transform. Wavelets were first introduced in 1987, as the foundation of a powerful new approach to signal processing, called multiresolution theory (Mallat 1999). Multiresolution theory incorporates and unifies techniques from a variety of disciplines, including subband coding from signal processing, quadrature mirror filtering from speech recognition, and pyramidal image processing (Gonzalez and Woods 2002). Formally, a multiresolution analysis (MRA) allows the representation of signals with their wavelet transform coefficients. The theory underlying MRA allows a systematical method for constructing (bi)orthogonal wavelets (Daubechies 1988) and leads to the fast discrete wavelet transform (DWT), also known as Mallat's pyramid algorithm (Qian 2002, Mallat 1989). In practice, the DWT has been applied to many different problems (Meyer 1990, Strang and Nguyen 1996, Daubechies 1992).

This Chapter aims to review wavelet transforms with a link to Mallat's pyramid algorithm and from the multiresolution point of view. It serves as the background theory for most of the work in this Thesis. Wavelet denoising will be considered later in this Chapter. In Chapter 7 and Chapters 10 to 12, wavelets will be used in the THz context for feature extraction, image segmentation and reconstruction.

6.1 Wavelet and multiresolution processing

Wavelet transforms are popular techniques suited to the analysis of very brief signals, especially signals with sudden and unpredictable changes that often carry the most interesting information (Qian 2002, Hubbard 1998). The wavelet transform technique is particularly well-suited to problems such as signal compression, feature extraction, image enhancement, and noise removal, which are especially related to biomedical applications. They complement the traditional Fourier based techniques in THz signal analysis in providing superior time-frequency localisation characteristics that are well-matched to the requirements for the short-duration T-ray pulse signals. Multiresolution analysis (MRA), also called multiscale approximation (MSA), is viewed as a most practical framework for understanding wavelet transforms and supplies the justification for discrete wavelet transform.

The first component to multiresolution analysis is function spaces. A multiresolution analysis considers the space of finite energy functions, $L^2(\mathbb{R})$. This space contains all real-world signals. It is assumed to consist of a sequence of indexed, nested subspaces $\{V_j\}_{j\in Z}$. The nested subspaces V_j satisfy the following conditions:

(i) Inclusion: each subspace V_j is contained in the next subspace. The function $L^2(\mathbb{R})$ in one subspace is in all the higher (finer) subspaces.

$$\cdots \subset V_{j+1} \subset V_j \subset \cdots \subset V_1 \subset V_0 \subset \cdots \subset L^2(\mathbb{R}) \tag{6.1}$$

(ii) Completeness: a function in the whole space comprises of the parts in each subspace.

- Upward completeness of the subspaces

$$\overline{\bigcup V_j} = L^2(\mathbb{R}) \tag{6.2}$$

  where $\overline{X}$ indicates the closed subspace of X

- Downward completeness

$$\bigcap V_j = 0 \tag{6.3}$$

(iii) Scale invariance

$$f(t) \in V_j \Rightarrow f(2t) \in V_{j-1} \tag{6.4}$$

(iv) Shift invariance

$$f(t) \in V_j \Rightarrow f(t-k) \in V_j \quad \forall k \in Z \tag{6.5}$$

(v) Basis-frame property: Multiresolution schemes require a basis for each space V_j. That is, there exists a scaling function $\phi(t) \in V_0$

$$\exists\, \phi(t) \in V_0\,, \quad \text{such that span } \{\phi_{0,k} | \forall k \in Z\} = V_0\,, \quad \text{where } \phi_{0,k} = \phi(t-k). \tag{6.6}$$

It can then be shown that, for $\forall j \in Z$, the set

$$\phi_{j,k}(t) = \phi(2^j t - k) \tag{6.7}$$

forms an orthonormal basis for V_j. It follows from orthogonality that

$$\int \phi_{j,k}(t)\phi_{j,k'}(t)dt = \delta_{k-k'}. \tag{6.8}$$

This means that for the same scale j, the scaling functions are orthonormal in time shifts. It is also an important feature for every scaling function $\phi(t)$, that there exists a set of coefficients $\{h(k)\}$, which satisfy the two-scale equation

$$\phi(t) = \sqrt{2}\sum_k h(k)\phi(2t-k). \tag{6.9}$$

It is built from translations of double-frequency copies of itself, $\phi(2t)$. This equation is sometimes called a multiresolution analysis equation. This set of coefficients $\{h(k)\}$ are called the scaling function coefficients, and can be thought of as a lowpass filter in discrete time.

Taking the Fourier transform of the both sides of the above equation, yields

$$\Phi(\omega) = \frac{1}{\sqrt{2}} H(\frac{\omega}{2})\Phi(\frac{\omega}{2}) \tag{6.10}$$

therefore, the scaling filter in the Fourier domain is

$$H(\omega) = \sqrt{2}\frac{\Phi(2\omega)}{\Phi(\omega)}. \tag{6.11}$$

According to orthonormal multiresolution analysis, the scaling function $\phi(t)$ in the frequency domain can also be computed by iterating equation (6.10)

$$\Phi(\omega) = \prod_{j=1}^{\infty} \frac{H(\omega/2^j)}{\sqrt{2}} \tag{6.12}$$

subject to the infinite product converging.

Define the approximation operation operator P_j on the functions $f(t) \in L^2(\mathbb{R})$ by

$$P_j\phi(t) = \sum_k \langle f(t), \phi_{j,k}\rangle \phi_{j,k}(t) \tag{6.13}$$

where $\langle f, \phi \rangle$ indicates the inner product between two integrable functions: $\langle f, \phi \rangle = \int_{-\infty}^{+\infty} f(t)\phi^*(t)\mathrm{d}t$, and $*$ indicates the complex conjugate.

Similarly, the detail operator Q_j on function $f(t) \in L^2(\mathbb{R})$ is defined by

$$Q_j\phi(t) = P_{j-1}f(t) - P_jf(t). \tag{6.14}$$

It is not difficult to show that, if $\{\phi_{j,k} | \forall j \in Z\}$ is an orthonormal basis for space V_j (Walnut 2001), then for all functions $f(t) \in L^2(\mathbb{R})$

$$\lim_{j \to -\infty} ||P_jf(t) - f(t)||_2 = 0 \tag{6.15}$$

and

$$\lim_{j \to \infty} ||P_jf(t)||_2 = 0 \tag{6.16}$$

where $||f||_2 = \sqrt{\langle f, f \rangle} = [\int_{-\infty}^{+\infty} f(t)^2]^{1/2}$.

Given a scaling function that meets the MRA requirements, a wavelet function ψ can be built from translations of $\phi(2t)$

$$\psi(t) = \sqrt{2}\sum_k g(k)\phi(2t - k). \tag{6.17}$$

This is a fundamental wavelet equation. The wavelet function coefficients are represented by the coefficients $\{g(k)\}$, which behaves as a highpass filter.

In the frequency domain, the wavelet equation can be rewritten

$$\Psi(\omega) = \frac{1}{\sqrt{2}} G(\frac{\omega}{2})\Phi(\frac{\omega}{2}). \tag{6.18}$$

Eq. (6.9) and Eq. (6.17) show that both the wavelet function and the scaling function are a weighted sum of scaling functions at the next finer scale. Let $\{V_j\}$ is a multiscale analysis with scaling function $\phi(t)$ and scaling filter $h(k)$, the wavelet filter $g(k)$ can be defined

$$g(k) = (-1)^k h(k)^*(1 - k) \tag{6.19}$$

where $h(k)^*$ indicates the complex conjugate of $h(k)$ in only the orthonormal case.

The Fourier transform of Eq. (6.19) is:

$$G(\omega) = -e^{-i\omega}H^*(\omega + \pi). \tag{6.20}$$

A consequence of the above structure of an orthonormal MRA is that it is possible to explicitly construct an orthonormal wavelet basis for $L^2(\mathbb{R})$. As the V_j is the real subspace of V_{j-1}, there exists the orthogonal complement subspace W_j, which satisfies

$$V_{j-1} = V_j \bigoplus W_j \tag{6.21}$$

where $\bigoplus$ denotes direct sum of vector spaces. Here, the vector spaces relate the scaling and wavelet function subspaces. Since the orthonormal complement of V_j in V_{j-1} is W_j, all members of V_j are orthogonal to the members of W_j. Thus,

$$\langle \phi_{j,k}(t), \psi_{j,k'}(t) \rangle = 0 \tag{6.22}$$

where $\langle \phi, \psi \rangle$ indicates the inner product between two integrable functions. The set $\psi_{j,k}(t)$ of wavelets is defined as

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k). \tag{6.23}$$

Several important facts that represent the key statements of multiresolution structure of the wavelet function are as follows.

(i) Completeness

$$\bigoplus_{j \in Z} W_j = L^2(\mathbb{R}) \tag{6.24}$$

(ii) Scale invariance

$$f(t) \in W_j \Rightarrow f(2t) \in W_{j-1} \tag{6.25}$$

(iii) Shift invariance

$$f(t) \in W_j \Rightarrow f(t-k) \in W_j \ \forall k \in Z. \tag{6.26}$$

The MRA structure presented in equations Eq. (6.1)-Eq. (6.9) can be applied to construct biorthogonal wavelets. A biorthogonal basis is composed from two mother wavelets, $\psi(t)$ and $\tilde{\psi}(t)$, which are duals of each other. The biorthogonality refers to the fact that dyadic dilations and translations of one mother wavelet is orthogonal to the other. This biorthogonal process can be interpreted from the multiresolution perspective. The analysis is specified by two hierarchical approximation subspaces:

$$\begin{aligned} &\cdots \subset V_{n+1} \subset V_n \subset \cdots \subset V_1 \subset V_0 \subset \cdots \subset L^2(\mathbb{R}) \\ &\cdots \subset \tilde{V}_{n+1} \subset \tilde{V}_n \subset \cdots \subset \tilde{V}_1 \subset \tilde{V}_0 \subset \cdots \subset L^2(\mathbb{R}). \end{aligned} \tag{6.27}$$

The orthonormal complementary properties hold for $W_j \perp \tilde{V}_j$ and $\tilde{W}_j \perp V_j$. This leads to

$$V_{j-1} = V_j \bigoplus \tilde{W}_j \text{ and } \tilde{V}_{j-1} = \tilde{V}_j \bigoplus W_j \tag{6.28}$$

and the following scaling and wavelet functions:

$$\begin{aligned} \phi(t) &= \sqrt{2}\sum h(k)\phi(2t-k) \\ \tilde{\phi}(t) &= \sqrt{2}\sum \tilde{h}(k)\tilde{\phi}(2t-k) \end{aligned} \tag{6.29}$$

and

$$\begin{aligned} \psi(t) &= \sqrt{2}\sum g(k)\phi(2t-k) \\ \tilde{\psi}(t) &= \sqrt{2}\sum \tilde{g}(k)\tilde{\phi}(2t-k) \end{aligned} \tag{6.30}$$

where

$$\begin{aligned} \tilde{g}_k &= (-1)^k h_{1-k} \\ g_k &= (-1)^k \tilde{h}_{1-k}. \end{aligned} \tag{6.31}$$

The scaling and wavelet functions are interrelated for the biorthogonal case:

$$\begin{aligned} \langle \tilde{\phi}(t-k), \phi(t-k) \rangle &= \delta_{k-1} \\ \langle \tilde{\psi}(t-k), \psi(t-k) \rangle &= \delta_{k-1} \end{aligned} \tag{6.32}$$

$$\begin{aligned} \langle \bar{\phi}(t-k), \psi(t-k) \rangle &= 0 \\ \langle \bar{\psi}(t-k), \phi(t-k) \rangle &= 0. \end{aligned} \tag{6.33}$$

Based on the foregoing, any function $x \in L^2(\mathbb{R})$ can be written as

$$\begin{aligned} f(t) &= \sum_j \sum_k \langle x, \tilde{\psi}_{j,k} \rangle \psi_{j,k}(t) \\ &= \sum_j \sum_k \langle x, \psi_{j,k} \rangle \tilde{\psi}_{j,k}(t). \end{aligned} \tag{6.34}$$

The equations in the frequency domain regarding biorthogonal wavelets are omitted for brevity.

To sum up, linear phase filters are one of the most important aspects achieved based on biorthogonality, which is vital in applications such as image compression, because phase distortions result in highly undesirable visual artifacts. Perfect reconstruction is another important property of orthogonal and biorthogonal wavelets, which will be discussed in Section 6.2.3. The filter relations for the orthonormal and biorthogonal wavelets as derived in this Section form the basis for the development of the Discrete Wavelet Transform (DWT).

6.2 Wavelet transforms in one dimension

The wavelet transforms are based on wavelets, which are small waves of varying frequency and limited duration. In subband coding, a transformation is computed by filtering and subsampling. The signal is separated approximately into frequency bands for efficient decoding.

6.2.1 Wavelet series expansions and discrete wavelet transforms

Wavelet series expansion of function $f(t) \in L^2(\mathbb{R})$ can be defined depending on wavelet $\psi(t)$ and scaling function $\phi(t)$. According to MRA, a set of time series function $f(t)$ can be approximated by projecting smooth information into multiple scaling subspaces and projecting the details information into wavelet subspaces, as in

$$f(t) = \sum_k c_{J_0}(k)\phi_{J_0,k} + \sum_k \sum_{j=0}^{J_0} d_j(k)\psi(j,k) \tag{6.35}$$

where J_0 is an arbitrary ending scale. The $c_{J_0}(k)$'s are called the approximation or scaling coefficients, and the $d_j(k)$'s are referred to as the detail or wavelet coefficients. This is because the Eq. (6.35) is composed of two parts, the first uses scaling functions to provide an approximation of $f(t)$ at scale J_0; for each higher scale $j \leq J_0$, the second part is a finer resolution function, a sum of wavelets, which is added to the approximation to provide increased detail. If the expansion functions form an orthogonal basis, the expansion coefficients of the time series function can be calculated as

$$c_{J_0}(k) = \langle f(t), \phi_{J_0,k}(t) \rangle = \int f(t)\phi^*_{J_0,k}(t)\mathrm{d}t \tag{6.36}$$

and

$$d_j(k) = \langle f(t), \psi_{j,k}(t) \rangle = \int f(t)\psi^*_{j,k}(t)\mathrm{d}t. \tag{6.37}$$

Compared to the wavelet series expansion presented above, which is to map a function of a continuous variable into a sequence of coefficients, the discrete wavelet transform (DWT) of $f(t)$ expands the function in a sequence of numbers. In this case, Eq. (6.35) can be rewritten as

$$f(t) = \sum_k c_{J_0}(k)\phi(2^{J_0}t - k) + \sum_k \sum_{j=0}^{J_0} d_j(k)\psi(2^j t - k) \tag{6.38}$$

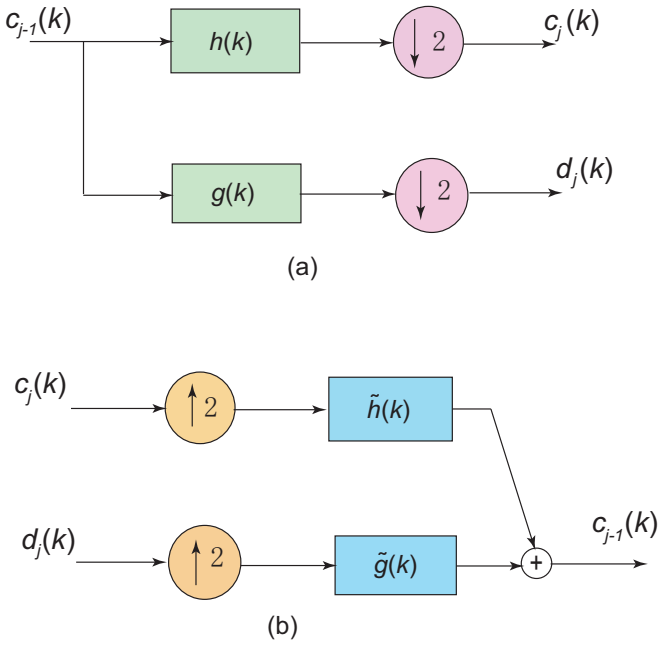


Figure 6.1. Block diagram for the analysis and synthesis stages of the Fast Wavelet Transform (FWT). The analysis (a) and synthesis (b) stages of Mallat's algorithm (FWT) are illustrated.

where $\phi(t)$'s are scaling functions and $\psi(t)$'s are wavelet functions; $f(t), \phi_{J_0,k}(t)$, and $\psi_{j,k}(t)$, all are functions of the discrete variable $t = 0, 1, 2, ..., M-1$; j is the index of scale in the range $\{0, J_0\}$, k is the time shift factor, and M is usually selected as a power of 2 (i.e. $M = 2^j$). The transform is composed from M coefficients, and the starting scale is 0, the ending scale is J_0. For the biorthogonal case, the functions ϕ and ψ should be replaced by their duals, $\tilde{\phi}$ and $\tilde{\psi}$.

Since biorthogonal bases consist of the dual pairs of ϕ and ψ, care should be taken that, if the expansion functions are the dual terms, the functions ϕ and ψ should be replaced in these equations by their dual functions $\tilde{\phi}$ and $\tilde{\psi}$.

6.2.2 The fast wavelet transforms

The fast wavelet transform (FWT) is a computationally efficient implementation of the discrete wavelet transform (DWT) that exploits a the relationship between the coefficients of the DWT at adjacent scales. It is also called Mallat's pyramid algorithm.

Given an orthogonal wavelet basis, the signal $f(t)$ can be represented by scaling series coefficients $\{c_{j_0}(k)\}$ and wavelet series coefficients $\{d_{j_0}(k)\}$, where J_0 is a desired

wavelet transform ending scale. Mallat's algorithm (or FWT), states the coefficients $\{c_j(k)\}_0^{j_0}$ and $\{d_j(k)\}_0^{J_0}$ can be calculated from coefficients at lower scales, by cascade filtering and subsampling with a pair of filters. The pair of filters consist of low-pass filter $H(\omega)$ and high-pass filter $G(\omega)$, which form a pair of mirror filters. The low-pass and high-pass filters $H(\omega)$ and $G(\omega)$ are the frequency responses of predefined filter coefficients $\{h(k)\}$ and $\{g(k)\}$, which characterise the scaling function and wavelet functions, respectively (Guo *et al.* 2001). The coarse and detailed coefficients at scale $j-1$ are calculated by applying the data at scale j and wavelet filtered banks as follows:

$$\begin{aligned} c_{j-1}(k) &= \sum_m c_j(m)\langle \phi_{j,m}, \phi_{j-1,k}\rangle = \sum_m h(m-2k)c_j(m) \\ d_{j-1}(k) &= \sum_m c_j(m)\langle \phi_{j,m}, \psi_{j-1,k}\rangle = \sum_m g(m-2k)c_j(m) \end{aligned} \tag{6.39}$$

where $m = 2k+n$; h and g correspond to low- and high-pass filters, respectively, and they satisfy Eq. (6.19).

The procedure is followed by discarding every second sample (downsampling by 2). The downsampling preserves the total amount of coefficients. At each level, the data, $c_{j-1}(k)$, is separated into two halves, or subbands: one contains the low frequency, or coarse information $(c_j(k))$, and the other half contains the high-frequency, or detailed information $(d_j(k))$. This is commonly known as the analysis (decomposition) stage, and is illustrated in Fig. 6.1(a). A full j-level FWT re-iterates equations Eq. (6.39) j times to yield a successive approximation to the original data $c_0(k)$. During this analysis stage, the higher level coefficients are calculated by the data at lower resolutions, and *vice versa*.

The inverse transform, can also be derived from the MRA structure. Reconstruction of signals from their scaling and wavelet coefficients can be realised by reversing Mallat's pyramid algorithm

$$\begin{aligned} c_j(k) &= \sum_m c_{j-1}(m)\langle \phi_{j-1,m}, \phi_{j,k}\rangle + \sum_m d_{j-1}(m)\langle \psi_{j-1,m}, \phi_{j,k}\rangle \\ &= \sum_m \tilde{h}(k-2m)c_{j-1}(m) + \sum_m \tilde{g}(k-2m)d_{j-1}(m) \end{aligned} \tag{6.40}$$

where $\tilde{h}$ and $\tilde{g}$ indicate the reconstruction low- and high-pass filters, respectively. For orthogonal wavelets, the reconstruction filers are simply time-reversed versions of the decomposition filters, that is, $h(k) = \tilde{h}(-k)$ and $g(k) = \tilde{g}(-k)$. In other words, $\tilde{h}$ and $\tilde{g}$ are often delayed versions of $h(-k)$ and $g(-k)$ to keep them causal. The net effect is a finite latency through the system (analysis plus synthesis).

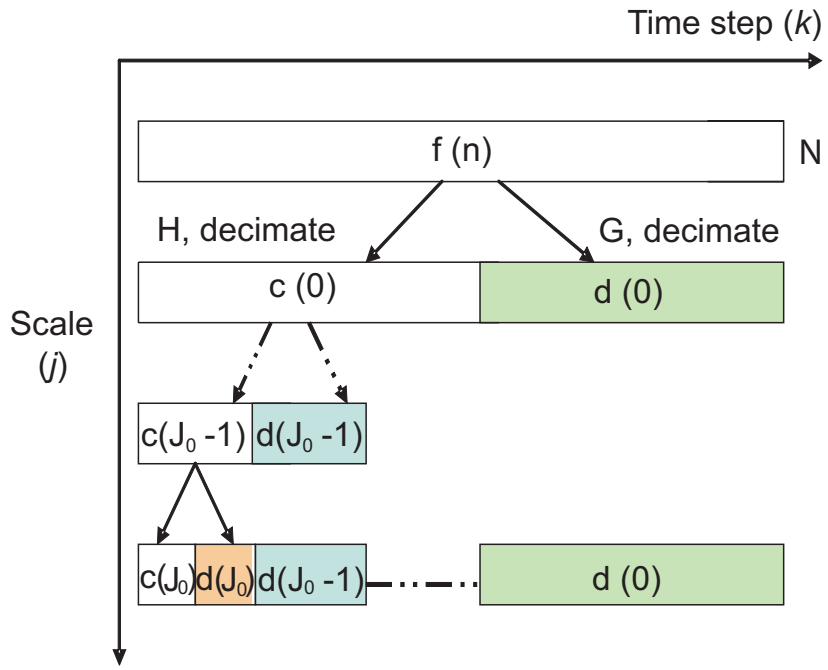


Figure 6.2. Fast Wavelet Transforms. The Fast Wavelet Transform (or Mallat's algorithm) is the calculation of a DWT using a digital filter bank followed by wavelet decimation (down sampling). This procedure operates on individual time frames at each wavelet scale.

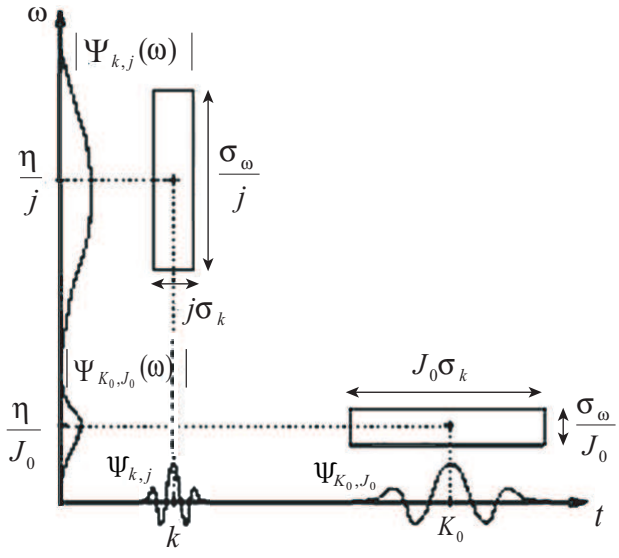


Figure 6.3. Time-frequency boxes of two wavelets $\psi_{j,k}$ and ψ_{J_0,K_0}. With the decreased scale j, the time support is reduced but the frequency spread increases and covers an interval that is shifted towards high frequencies. After Mallat (1999).

The one step reconstruction represented in Eq. (6.40) can be iteratively applied to multiple levels of wavelet reconstruction, until the original signal is recovered. Each reconstruction stage adds the detail subbands $d_j(k)$ back to the coarse approximation subband $c_j(k)$ to form the finer approximation $c_{j+1}(k)$. Each synthesis (reconstruction) stage needs interpolation of each subband by 2, followed by filtering with the synthesis filters, $\tilde{h}(k)$ and $\tilde{g}(k)$, then summing up the two sets of resultant coefficients. This synthesis processing is illustrated in Fig. 6.1(b). These operations are repeated over all levels to reconstruct the original signal $c_0(k)$. The entire analysis process in a J-level FWT is illustrated in Fig. 6.2. The number of computations is halved with an increase in the level of the transform, as a direct consequence of the reduced data from downsampling. This reduction in complexity with levels leads to the high efficiency of the algorithm. The fast wavelet transform of a length $M = 2^{J_0}$ sequence has an overall computational complexity of $O(M\log M)$. Since the analysis and synthesis stages employ two filters (low- and high-pass), the FWT system is commonly called a 2-channel filter bank in the engineering literature. As for the practical application, the sampled T-ray transients are a function of discrete time, to which the discrete wavelet transform can be applied. To realize the DWT, simple digital filter banks are utilised in a recursive structure to calculate wavelet transform coefficients of T-ray signals. As a result, a representation of the original T-ray pulsed signals consisting of different scales or resolutions at different times (positions) is achieved. Fig. 6.3 shows time-frequency boxes of two wavelets $\psi_{j,k}$ and ψ_{J_0,K_0}. In the time-frequency plane, a wave atom $\psi_{k,j}$ is symbolically represented by a rectangle centered at $(k, \eta/j)$. The energy of $\Psi_{k,j}$ (Fourier transform of $\psi_{k,j}$) is concentrated over a positive frequency interval of σ_ω/j centered at η/j. The time and frequency spread are respectively proportional to j and $1/j$. When j varies, the height and width of the rectangle change but its area remains constant.

6.2.3 Perfect reconstruction of 2-channel filter bank

A filter bank is a set of filters that mainly link to sampling operations. In a two-channel filter bank, the analysis filters are lowpass $H(z)$ and highpass $G(z)$, where $H(z)$ and $G(z)$ in the frequency domain are z-transforms of $h(k)$ and $g(k)$ in the time domain. The synthesis bank for signal reconstruction performs the inverse procedure to analysis, which means lowpass $\tilde{H}(z)$ and highpass $\tilde{G}(z)$ filtering (z-transforms of $\tilde{h}(k)$ and $\tilde{g}(k)$), and interpolating. In order to recover the discarded information by the downsampling in the analysis bank, the synthesis filter bank must be specifically adapt to the

analysis filters for perfect reconstruction. This means that the filter bank is biorthogonal. An orthonormal filter bank that the synthesis is the transpose of the analysis, is a specific case of the more general biorthogonal filter bank.

There are two conditions which need to be imposed to achieve perfect reconstruction. The first condition is to remove aliasing in the reconstructed signal, which leads to as in

$$\tilde{H}(z)H(-z) + \tilde{G}(z)G(-z) = 0. \tag{6.41}$$

The second condition is to eliminate amplitude distortion, which leads to

$$\tilde{H}(z)H(z) + \tilde{G}(z)G(z) = 2 \tag{6.42}$$

where $\tilde{G}(z)$ is a function of $H(z)$ and $G(z)$ is a function of $\tilde{H}(z)$. Note that for Eq. (6.41), if we define $S(z) = H(z)\tilde{H}(z)$, then this equation becomes $S(z) + S(-z) = 2$. This implies that $s(k)$ is a half-band filter, with zeros for all its even coefficients excluding $s(0) = 1$. The odd coefficients are then free design variables. In practice, the wavelet filters are usually derived by first designing a particular half-band filter $s(k)$. The analysis and synthesis low-pass filters are then obtained by applying spectral factorisation to $s(k)$.

The orthonormality condition for perfect reconstruction filter banks is defined in the time domain as

$$\langle g_i(n), g_i(n+2m)\rangle = \delta(i-j)\delta(m), \;\; i,j = \{0,1\}. \tag{6.43}$$

The biorthonormality condition is different from orthonormality condition, which is given in the time domain

$$\langle h_i(2n-k), g_j(k)\rangle = \delta(i-j)\delta(n), \;\; i,j = \{0,1\}. \tag{6.44}$$

It is useful to note that the biorthogonality condition holds for all two-band, real-coefficient perfect reconstruction filter banks.

It can be observed via the previous analysis that despite their central roles in the development of wavelet theory, the FWT algorithm does not apply the scaling and wavelet function for time series analysis. Instead, it is only necessary to explicitly construct the analysis and reconstruction filters $\{\, h(k), g(k), \tilde{h}(k), \tilde{g}(k) \,\}$ in order to compute the wavelet transform coefficients.

The drawback of the FWT is the different filter lengths in the decomposition part, caused by an odd sample shift in the input signal in the downsampling operations. It means that the signal spectrum is unevenly distributed over the low-band and high-band segments. In many practical situations, this shift variance is an undesirable aspect of the FWT. Researchers have spent significant efforts in overcoming this problem, leading to a variety of different approaches (Ng 2003). These will not be discussed further in this Thesis.

6.3 Two dimensional discrete wavelet transforms

Wavelet transforms play an important role in many image processing algorithms. In two dimensions, the discrete version of a wavelet transform can be realised by a 2D scaling function, $\phi(x,y)$, and three 2D wavelets, $\psi^H(x,y)$, $\psi^V(x,y)$, and $\psi^D(x,y)$, which are calculated by taking the 1D wavelet transform along the rows of $f(x,y)$ and the resulting columns (Gonzalez and Woods 2002). The 2D scaling function and 2D wavelet functions satisfy the following equations:

$$\phi(x,y) = \phi(x)\phi(y) \tag{6.45}$$
$$\psi^H(x,y) = \phi(x)\psi(y) \tag{6.46}$$
$$\psi^V(x,y) = \psi(x)\phi(y) \tag{6.47}$$
$$\psi^D(x,y) = \psi(x)\psi(y) \tag{6.48}$$

where $\phi(t)$ and $\psi(t)$ are the one-dimensional scaling and wavelet functions, respectively. There exists the sequences $c_{j-1}(k,l)$, $d^H_{j-1}(k,l)$, $d^V_{j-1}(k,l)$, and $d^D_{j-1}(k,l)$ given by the following equations:

$$\begin{cases} c_{j-1}(k,l) = \langle f(x,y), \phi_j(k,l)\rangle \\ d^H_{j-1}(k,l) = \langle f(x,y), \psi^H_j(k,l)\rangle \\ d^V_{j-1}(k,l) = \langle f(x,y), \psi^V_j(k,l)\rangle \\ d^D_{j-1}(k,l) = \langle f(x,y), \psi^D_j(k,l)\rangle. \end{cases} \tag{6.49}$$

The coefficients d^H_{j-1}, d^V_{j-1} and d^D_{j-1} correspond to horizontal, vertical, and diagonal high frequency information, respectively, while c_{j-1} corresponds to coefficients representing low frequency information (Mallat 1999).

The 2D Wavelet Transform algorithm can be derived from Eq. (6.49). Simply, the 2D DWT can be calculated by cascaded 1D DWTs along the rows and columns of the image

(Mallat 1999). This process can be repeated a number of times to yield successively lower resolution images. Mathematically, the process of taking one level of the 2D wavelet transform is:

$$\begin{cases} c_{j-1}(k,l) = \sum_{m,n} h(2k-m)h(2l-n)c_j(m,n) \\ d^H_{j-1}(k,l) = \sum_{m,n} h(2k-m)g(2l-n)c_j(m,n) \\ d^V_{j-1}(k,l) = \sum_{m,n} g(2k-m)h(2l-n)c_j(m,n) \\ d^D_{j-1}(k,l) = \sum_{m,n} g(2k-m)g(2l-n)c_j(m,n). \end{cases} \tag{6.50}$$

Let h and g denote a pair of linear phase low- and high-pass wavelet filters and $\tilde{h}$, $\tilde{g}$ denote the corresponding reconstruction filters. The discrete approximation at resolution 2^j can be obtained by combining the details and approximation at resolution 2^{j-1} using reconstructed wavelet filters:

$$\begin{aligned} c_j(k,l) \;=\; & \sum_{m,n} \tilde{h}(k-2m)\tilde{h}(l-2n)c_{j-1}(m,n) \\ & +\tilde{h}(k-2m)\tilde{g}(l-2n)d^H_{j-1}(m,n) \\ & +\tilde{g}(k-2m)\tilde{h}(l-2n)d^V_{j-1}(m,n) \\ & +\tilde{g}(k-2m)\tilde{g}(l-2n)d^D_{j-1}(m,n). \end{aligned} \tag{6.51}$$

Fig. 6.4 shows the block diagram of the 2D wavelet transform procedure. The block diagram is for one step of the 2D discrete wavelet transform (2D DWT) and can be implemented for recovery of an approximation at resolution of 1 step.

6.4 Discrete wavelet packet transforms

Discrete wavelet packet transforms are the link between multiresolution approximations and discrete wavelets (Mallat 1999). The detailed subspace W_j is spanned can be calculated by the discrete wavelet basis $\psi_j(2^j t - k)$. The time and frequency spread (window) of the discrete wavelet basis $\psi_j(2^j t - k)$ are respectively proportional to j and $1/j$. With the increase of the scale j, the width of the time-frequency window of the wavelet basis has decreased, while its height is increased. This indicates that wavelet basis has contracted in its space localisation, but with an expansion in the frquency localisation. In order to overcome this problem, we need to decompose the W_j further. Discrete wavelet packet transforms (DWPTs) apply the discrete wavelet transform step to both the lower resolution and detail subspaces. This procedure is

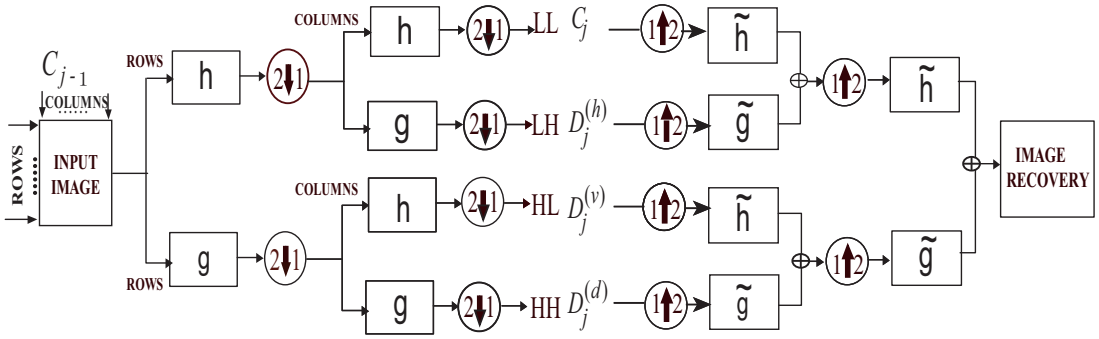


Figure 6.4. Illustration of the 2D discrete wavelet transform procedure. The 2D DWT can be realised via digital filters and downsampling the T-ray image. A 2D scaling function, $\phi(x,y)$ and three 2D wavelets, $\psi^H(x,y), \psi^V(x,y)$, and $\psi^D(x,y)$, are calculated via taking the 1D fast wavelet transform of the rows of $f(x,y)$ and the resulting columns. The discrete approximation at resolution 2^j can be obtained by combination of the details and approximation at resolution 2^{j-1} using reconstructed wavelet filters. Here, h & g and $\tilde{h}$ & $\tilde{g}$ indicate the low-pass and high-pass wavelet filters and reconstruction filters, respectively. The down and up arrows indicate the downsampling and upsampling procedures.

fulfilled by dividing the orthogonal basis $\{\phi_{j-1}(2^j t-k)\}_{k\in z}$ into two new orthogonal bases: $\{\phi_j(2^j t-k)\}_{k\in z}$ of V_j and $\{\psi_j(2^j t-k)\}_{k\in z}$ of W_j, where

$$\phi(t)=\sqrt{2}\sum_{k\in z} h(k)\phi(2t-k) \tag{6.52}$$

$$\psi(t)=\sqrt{2}\sum_{k\in z} g(k)\phi(2t-k). \tag{6.53}$$

For the analysis of the wavelet packet function, new signs $\mu_0(t)$ and $\mu_1(t)$ are used instead of $\phi(t)$ and $\psi(t)$,

$$\mu_0(t)=\sqrt{2}\sum_{k\in z} h(k)\mu_0(2t-k) \tag{6.54}$$

$$\mu_1(t)=\sqrt{2}\sum_{k\in z} g(k)\mu_0(2t-k). \tag{6.55}$$

For a fixed decomposing scale, the wavelet packet function can be defined by $\mu_0(t)$, $\mu_1(t)$, h, and g,

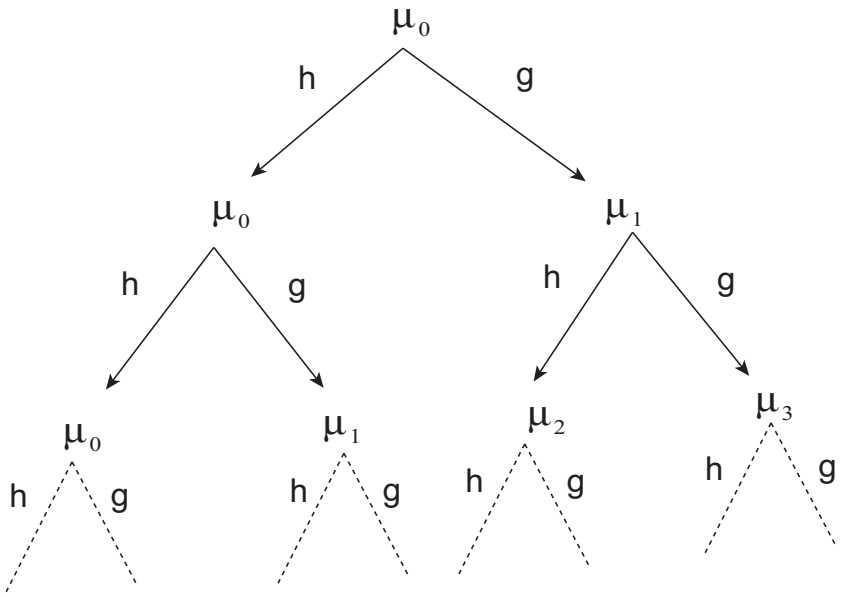


Figure 6.5. The procedure of discrete wavelet packet transform. Time-localization tree after wavelet package decomposition.

$$\mu_{2n}(t) = \sqrt{2}\sum_{k} h(k)\mu_n(2t-k) \tag{6.56}$$

$$\mu_{2n+1}(t) = \sqrt{2}\sum_{k} g(k)\mu_n(2t-k) \tag{6.57}$$

where $n = 0, 1, 2...$; g and h are a pair of conjugate mirror filters. The function μ_n can be iteratively calculated based on the Eqs. (6.56) and (6.57). It can be derived according to the orthogonal scaling function $\mu_0 = \phi$.

This discrete wavelet packet can be represented by a time-localization tree. The procedure of DWPT is shown in Fig. 6.5.

Several important properties of discrete wavelet packet transforms are: (i) For $\forall n \geq 0$, if the binary system of n satisfies the equation

$$n = \sum_{j=1}^{\infty} \varepsilon_j 2^j, \;\; \varepsilon_j = 0 \;\text{ or }\; 1 \tag{6.58}$$

the Fourier transforms of wavelet packet μ_n

$$\hat{\mu}_n(\omega) = \Pi_{j=1}^{\infty} m_{\varepsilon_j}\left(\frac{\omega}{2^j}\right) \tag{6.59}$$

where

$$m_0(\omega) = \frac{1}{\sqrt{2}} \sum_k h(k) e^{-ik\omega}$$
$$m_1(\omega) = \frac{1}{\sqrt{2}} \sum_k g(k) e^{-ik\omega} \tag{6.60}$$

where $\sum$ and $\prod$ present the sum and product of finite sequences.

(ii) For $\forall n \geq 0$, μ_n satisfies translation orthogonality

$$\langle \mu_n(t-j), \mu_n(t-k) \rangle = \delta_{j,k}, \quad j,k \in z. \tag{6.61}$$

(iii) For $\forall n \geq 0$, μ_{2n} and μ_{2n+1} satisfies translation orthogonality

$$\langle \mu_{2n}(t-j), \mu_{2n+1}(t-k) \rangle = \delta_{j,k}, \quad j,k \in z. \tag{6.62}$$

(iv) The wavelet packet function $\{\mu_n(t-k) |\ n \in z_+ \text{ and } k \in z\}$ is composed of a group of basis of $L^2(\mathbb{R})$. Let

$$\begin{aligned} U_j^0 &= V_j \\ U_j^1 &= W_j, \quad j \in z \end{aligned} \tag{6.63}$$

and we can rewrite

$$V_{j-1} = V_j \bigoplus W_j, \quad j \in z \tag{6.64}$$

to

$$U_{j-1}^0 = U_j^0 \bigoplus U_j^1, \quad j \in z. \tag{6.65}$$

For $\forall n \in z_+$, we get

$$U_{j-1}^n = U_j^{2n} \bigoplus U_j^{2n+1}, \quad j \in z \tag{6.66}$$

where $U_j^{2n} \perp U_j^{2n+1}$.

Now, we calculate discrete wavelet packet coefficients of the time series $f(t)$ via the pair of mirror filters of DWT, with the number of wavelet decomposition steps of $N = 2^L$. The aim of the one step in the discrete wavelet packet transform is to calculate a higher resolution subspace U_j^{2n} and a detailed subspace U_j^{2n+1} by the subspace ${U_{j-1}}^n$ of U_0^0, $j = 0, 1, 2...L-1, L$. The orthonormal basis of subspace U_{j-1}^n can be constructed as in

$$2^{(j-1)/2} \mu_n(2^{j-1}t - k), \ \ k \in z. \tag{6.67}$$

So that the discrete wavelet packet coefficients of $f(t)$ for projection onto subspace U_{j-1}^n can be calculated

$$d_{j-1}^n(k) = \langle f(t), 2^{(j-1)/2}\mu_n(2^{j-1}t - k)\rangle. \tag{6.68}$$

The root node at $j = 0$ can be viewed as the approximation of the sampled signal $f(t)$, which can be represented by the discrete wavelet packet coefficients, noted as $d_0^0(k)$.

Similarly, the discrete wavelet packet coefficients for projections onto U_j^{2n} and U_j^{2n+1} can be noted as d_j^{2n} and d_j^{2n+1}, respectively. Based on the discrete wavelet packet definition, Eq. (6.66), we can achieve the discrete wavelet packet decomposition and reconstruction algorithms. At the decomposition:

$$\begin{cases} d_{j-1}^{2n}(k) = \sum_{l\in z} h(l-2k)d_j^n(l) \\ d_{j-1}^{2n+1}(k) = \sum_{l\in z} g(l-2k)d_j^n(l). \end{cases}$$

At the reconstruction:

$$d_j^n(k) = \sum_{l\in z}(\tilde{h}(k-2l)d_{j-1}^{2n}(l) + \tilde{g}(k-2l)d_{j-1}^{2n+1}(l)). \tag{6.69}$$

Here $\tilde{h}$ and $\tilde{g}$ are the reconstruction low- and high-pass filters, respectively. The discrete wavelet package coefficients d_{j-1}^{2n} and d_{j-1}^{2n+1} are the subsampled convolutions of d_j^n with h and g, and $d_{j-1}^n(k)$ satisfies

$$d_{j-1}^n(k) = \langle f(t), \psi_{j-1}^n(2^{j-1}t - k)\rangle. \tag{6.70}$$

Here, j is the node depth in the wavelet packet binary tree and n denotes the number of the note at the same node depth. Two scales of the wavelet packet tree in Fig. 6.6 results in almost twice the number of decompositions and reconstructions that are available from the two-scale DWT.

Different wavelet packet bases show different capabilities to achieve time-frequency localisation and represent different properties of signals. It is important to select a best wavelet packet basis from the dictionary of bases. A few questions arise: how to select the best wavelet packet basis? What is the criteria to evaluate the good adaptability of a wavelet packet basis? How to make a quick search within the dictionary for the best basis?

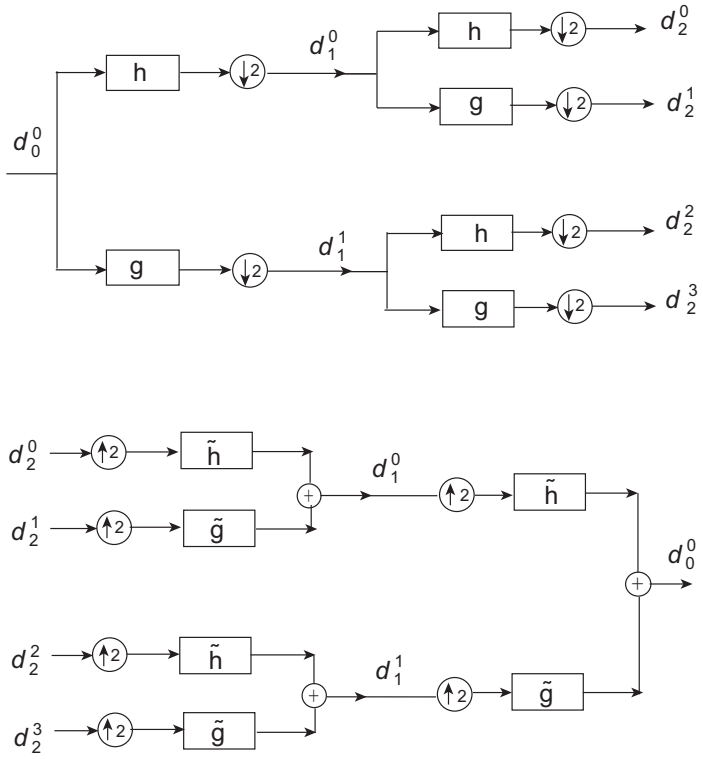


Figure 6.6. The procedure of discrete wavelet packet transform. Illustration of the two-level discrete wavelet packet decompositions.

Wavelet packet bases are large families of orthogonal bases that include different types of time-frequency atoms. The best wavelet packet basis decomposes signals over atoms that are adapted to the signal's time-frequency features. To evaluate the best basis, first we perform wavelet series spread of the signal $f(t)$ depending on an orthonormal wavelet packet basis. It turns out to be wavelet packet coefficient series $E = \{E_k\}$ corresponding to $f(t)$. We define an information cost function W on the series, and it should satisfy two conditions:

(i) The additivity

$$W(\{E_k\}) = \sum_{k \in z} W(E_k), \quad W(0) = 0. \tag{6.71}$$

(ii) The value of the information cost function W should reflect the concentration degree of wavelet packet coefficients. It can be explained as follows:

- If the energy of the coefficient series $\{E_k\}$ is concentrated to a few coefficients and the absolute value of a large number of the coefficients is small enough to be ignored, the corresponding basis is considered to be better, as the value of the information cost function should be smaller.

- If the energy of coefficient series $\{E_k\}$ is distributed uniformly, the correspondent basis is not desirable, and the value of the information cost function W should be larger. Generally, the wavelet packet basis B of $L^2(\mathbb{R})$ for effectively approximating a signal $f(t)$ should be optimised by minimising the knowledge of an information cost function.

The following are several frequently used information cost functions:
(i) Calculation of the number of the wavelet packet coefficients with magnitude larger than a threshold.
We predefine a threshold T to calculate the number of the elements in the series $\{E_k\}$ with an absolute value over the threshold T,

$$W(\{E_k\}) = \begin{cases} 1 & |E_k| \geq T \\ 0 & |E_k| < T. \end{cases} \tag{6.72}$$

(ii) The concentration degree of the $\mathbf{l^p}$ norm of the wavelet packet coefficients.
For $\forall 0 < p < 2$, we define $W(E_k)^p = |E_k|^p$, therefore, $W(E) = \sum_{k\in z} |E_k|^p = ||E||^p$. Usually, the L^p norm is defined as $||x||_p = (\sum x_i^p)^{\frac{1}{p}}$.

(iii) Logarithmic entropy.
For $\forall E = \{E_k\}$, we define $W(E) = \sum_{k\in z} \log|E_k|^2$, and $\log 0 = 0$.

(iv) Information entropy.
For $\forall E = \{E_k\}$, we define $W(E) = -\sum_{k\in z} |E_k|^2 \log|E_k|^2$, and $\log 0 = 0$.

A basis pursuit is computationally expensive because it minimises a global cost function over all dictionary vectors. Here, we introduce a dynamic programming approach. Further discussion on the basis pursuit, i.e. matching pursuit, can be found in Mallat (1999).

The dynamic programming calculation considers a discrete wavelet packet decomposition in space $V_0 = U_0$. This decomposition can be presented using a dual tree, as shown in Fig. 6.6. The wavelet expansion of $f(t)$ produces wavelet packet coefficients, which are used to calculate the value of information cost function at each node of the dual tree. Finally we trace the best basis node by node from bottom to top of the tree. The operation is as follows:

(1) We fix the maximum depth d allowed for the tree and initialize the search with a complete tree with that depth.

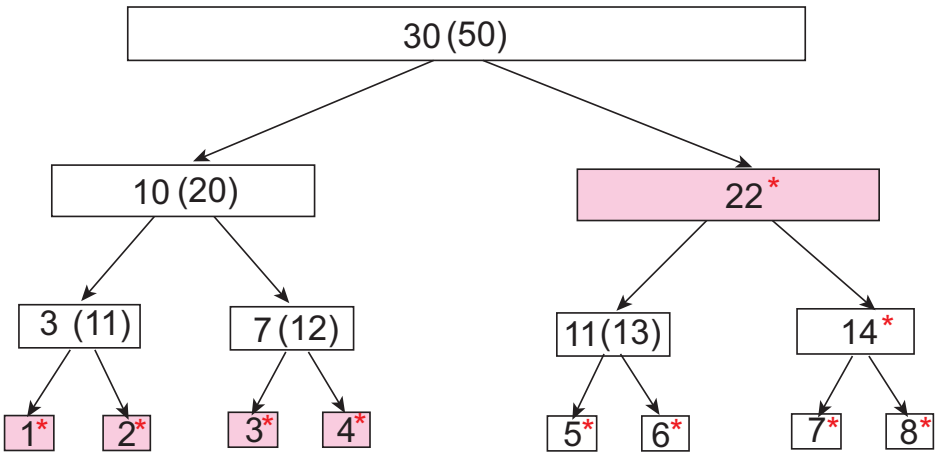


Figure 6.7. A quick searching of the best bases. The nodes with pink shade indicate the best wavelet packet bases. The number at the bottom row (the maximum depth) are supposed to be given, which are indicated by 1-8 digital numbers. We compare the value of cost function at the parent node W_f with the sum of the cost functions at its two child nodes W_b. If $W_f \leq W_b$, we label the parent node with '$*$', viewed as a leaf node. If $W_f > W_b$, we bracket the value W_f and put the value W_b beside the bracket. In this case, the corresponding nodes are viewed as non-leaf. We consider only the value outside the bracket and repeat the first and second step till the top layer of the tree.

(2) All nodes are candidates to be leaf nodes. For each node (i, j), leaf or non-leaf, we compare the value of cost function W_f at the parent node with the sum of the cost functions W_b at its two child nodes.

(3) Initialize the costs at level d, the deepest level of the tree. If one node of this level is kept in the tree, it will be necessarily a leaf node. This step is to first label the bottom node via the notation '$*$'.

(4) If $W_f \leq W_b$, we label the parent node with '$*$', to emphasize that the cost value at the parent node is not changed, which is viewed as a leaf node. If $W_f > W_b$, we bracket the value W_f and put the value W_b beside the bracket. In this case, the corresponding nodes are viewed as non-leaf nodes.

(5) We consider only the value outside the bracket and repeat the first and second step till the top layer of the tree.

(6) The final step is to select the labelled nodes that are close to the root. The procedure is performed from top to bottom of the whole tree (any subnodes, including the nodes, with and without a $*$, are not considered again). These selected nodes cover the space $V_0 = U_0^0$ without overlapping, all of which are viewed as the nodes corresponding to

the best bases. The Fig. 6.7 illustrates the linear searching procedure regarding the best wavelet packet bases.

6.5 Wavelet denoising for THz-TDS pulses via the heuristic SURE threshold

Wavelet denoising is an important application of wavelet transforms with respect to multiresolution techniques. In principle, wavelet-based techniques are very well suited to studies of non-stationary time-domain data sets, highlighting the variability of features at different time-frequency scales (Mallat 1999, Donoho 1995). The main concern with current T-ray TPI measurements is corruption by different types of noise which limit the practical usefulness of this mode of imaging. One of the main problems is the noise of the T-ray pulse as it propagates through the optical system. As a result, the received T-ray signals are strongly dependent on the acquisition conditions, and similar samples may produce variations in the measured signals in different portions of the image (Ferguson and Abbott 2001a). In order to keep image acquisition within realistic timescales, a short integration time per pixel is adopted as common practice, and this makes the choice of the THz-transient de-noising process attractive for improved signal conditioning.

The use of perfect reconstruction quadrature mirror filter banks has been extensively discussed by Daubechies (1992) for the purpose of de-noising and generating bases of compact support. The work of Vetterli and Kovacevic (1995) as well as that of Strang and Nguyen (1996) further complement the above, elaborating more on sub-band coding, Sherlock and Monro (1998) discuss how to apply FIR filters of arbitrary length to describe the space of orthonormal wavelets, further parameterizing the wavelet coefficients at each decomposed level, and Tuqun and Vaidyanathan (2000) propose a state-space approach to the design of globally optimal FIR energy compaction filters.

Since, in our work, there is no requirement for adopting an algorithm with a perfect reconstruction property, as our ultimate goal is feature extraction and classification, our constrains are more relaxed compared to those used in filtering or signal compression applications.

Divine and Godtliebsen (2007) suggest that for feature exploration purposes, it is possible to assume stationarity over some time interval and smooth the wavelet spectrum

along the time axis using an Auto Regressive (AR) model. Recently, Paiva and Galvão (2006) also discuss a wavelet packet decomposition tree algorithm that establishes frequency bands where sub-band models are created. Both approaches propose the modeling of the approximation and detail wavelet coefficients in order to further extract statistically significant features and a similar approach is adopted in our work. A typical de-noising procedure consists of decomposing the original signal using DWPT or DWT (Mallat 1999, Daubechies 1992, Jensen and la Cour-Harbo 2001, Percival and Walden 2000), thresholding the detail coefficients, and reconstructing the signal by applying the appropriate inverse transform (IDWT or IDWPT respectively). For the de-noising of femtosecond THz transients, a three-level decomposition is usually sufficient (Hadjiloucas *et al.* 2004) and unnecessary computational load associated with more decomposition levels can be avoided. The Stein's Unbiased Risk Estimate (SURE) and the 'heuristic' SURE methods (Donoho 1995) are used separately to estimate the soft threshold parameter (λ^S) for classification experiments.

In our T-ray project, each time domain measurement corresponding to data from a single pixel is represented by a data vector $\mathbf{x}$ of length L_x, where the n^{th} element of $\mathbf{x}$, denoted by x_n, represents the measured signal at the n^{th} sampling instant. The filter bank transform can be regarded as a change in variables from $\mathcal{R}^{L_x}$ to $\mathcal{R}^{L_x}$ performed according to the following operation

$$w_m = \sum_{n=0}^{L_x-1} x_n v_m(n), m = 0, 1, ..., L_x - 1 \tag{6.73}$$

where w_m is a transformed variable and $v_m(n) \in \mathcal{R}$ is a transform weight. It proves convenience to write the transform in matrix form as:

$$\mathrm{w}_{1\times L_x} = \mathbf{x}_{1\times L_x}\mathbf{V}_{L_x\times L_x} \tag{6.74}$$

where $\mathbf{x} = [x_i : i = 0, 1, ..., L_x - 1]$ is the row vector of original variables, $\mathbf{w}$ is the row vector of new (transformed) variables and $\mathbf{V}$ is the matrix of weights. Choosing $\mathbf{V}$ to be unitary (that is, $\mathbf{V}^{\mathrm{T}}\mathbf{V} = \mathbf{I}$), the transform is said to be orthogonal and it therefore consists of a simple rotation in the coordinate axes (with the new axes directions determined by the columns of $\mathbf{V}$). As described in wavelet literature, the discrete wavelet transform can be calculated in a fast manner by using a filter bank structure. In this filter bank, the low-pass filtering result undergoes successive filtering iterations with the number of iterations j chosen by the analyst. The final result of the decomposition of data vector $\mathbf{x}$ is a vector resulting from the concatenation of row vectors c_j and d_j

(termed respectively approximation and detail coefficients at the j^{th} scale level) in the following manner:

$$\mathbf{w} = [\mathrm{c}_j|\mathrm{d}_j|\mathrm{d}_{j-1}|...|\mathrm{d}_1] \tag{6.75}$$

with coefficients in larger scales (e.g. d_j,d_{j-1},d_{j-1},...) associated with broad features in the data vector, and coefficients in smaller scales (e.g. d_1,d_2,d_3,...) associated with narrower features such as sharp peaks. Let $h[0], h[1], ..., h[2N-1]$ and $g[0], g[1], ..., g[2N-1]$ be the impulse responses of the low-pass and high-pass filters H and G, respectively. Assuming that filtering is carried out by circular convolution, the procedure for generating the approximation coefficients from the data vector $\mathbf{x}$ consists of flipping the filtering sequence and moving it alongside the data vector. For each position of the filtering sequence regarding the data vector, the scalar product of the two is calculated. Dyadic down-sampling is then performed to generate coefficients $c[i]$. The detail coefficients $d[i]$ are obtained in a similar manner by using the high-pass filtering sequence. Filtering in the wavelet domain consists of changing some of the above elements of $\mathbf{w}$ by applying soft thresholding so that a new vector $\mathbf{w}_f$ is produced and then applying the inverse transform. A soft threshold operation with threshold λ^s is employed:

$$\mathbf{w}_f(\mathbf{x}, \lambda^s) = \mathrm{sgn}(\mathbf{x}) \max(0, |\mathbf{x}| - \lambda^s). \tag{6.76}$$

Soft thresholding produces better continuity (Donoho 1995) than hard thresholding, while at the same time, it provides shrinkage of the wavelet coefficients dominated by noise. The nonlinear shrinking of coefficients in the wavelet domain is a nonparametric method. We adopt the heuristic Stein's Unbiased Risk Estimate (SURE) (*Florida State University* 2005) algorithm, which is estimated adaptively at each wavelet decomposition level j. This operates on detail coefficients d_j. Each detail coefficient d_j is composed of elements of the signal d_f as well as superimposed noise series $\{\eta_j\}$, so that $\mathrm{d}_j = \mathrm{d}_{f_j} + \{\eta_j\}$.

The SURE approach assumes data series $\{\mu_j\} = (\mu_j[i] : i = 1, ..., n)$ of length $n = L_x/(2 \times j)$, where $d_j[i] \sim \eta(\mu_j[i], 1)$ is multivariate normal observations with that mean vector. The soft threshold estimator $\{\hat{\mu}_j\}$ of $\{\mu_j\}$ is given from $\hat{\mu} = \hat{\mu}(\mathrm{d}_j) = \mathrm{d}_j + \mathrm{g}(\mathrm{d}_j)$ with $\mathrm{g}_j = (g_j[i])_{i=1}^n$. Assuming that $\mathrm{g}_j(\mathrm{d}_j)$ is weakly differentiable, Stein's method calculates the energy loss $||\hat{\mu} - \mu||$ in an unbiased fashion using the following

$$\mathbb{E} \parallel \{\hat{\mu}_j\}(\mathrm{d}_j) - \{\mu_j\} \parallel^2 = n + \mathbb{E}\{||\mathrm{g}_j(\mathrm{d}_j)||^2 + 2\nabla \cdot \mathrm{g}_j(\mathrm{d}_j)\} \tag{6.77}$$

where $\nabla \cdot \mathrm{g}_j \equiv \sum_i \frac{\partial}{\partial \mathrm{d}_j[i]} g_j[i]\ \mathbb{E} \parallel \{\hat{\mu}_j\}^{(\lambda_j^s)}(\mathrm{d}_j) - \{\mu_j\} \parallel^2$.

By applying Stein's results (Donoho 1995, Johnstone and Donoho 1995) at the j^{th} decomposition level, we have

$$\mathrm{SURE}_j(\lambda_j^s; \mathrm{d}_j) = n - 2 \cdot \#\{i : |\mathrm{d}_j[i]| < \lambda_j^s\} + \sum_{i=1}^{n} \min(|\mathrm{d}_j[i]|, \lambda_j^s)^2 \tag{6.78}$$

where $\hat{\mu}_j$ denotes the soft threshold estimator $\hat{\mu}_j[i]^{(\lambda_j^s)}(\mathrm{d}_j) = \eta_{\lambda_j}^s(\mathrm{d}_j[i])$. Thus the threshold λ^s that minimizes $\mathrm{SURE}_j(\lambda_j^s; \mathrm{d}_j)$ is obtained. Here, # means the number of the coefficients that are smaller than the threshold.

The denoised T-ray signal in the wavelet domain can be re-written as

$$\mathbf{w}_f = [\hat{c}_j | \hat{\mathbf{d}}_{f_j} | \hat{\mathbf{d}}_{f_{j-1}} | ... | \hat{\mathbf{d}}_{f_1}]. \tag{6.79}$$

The equations derived above will be further taken advantage of in the next Chapter for feature extraction and classification application. For wavelet packet SURE denoising, we calculate the SURE threshold for each of WP coefficients, and then apply the calculated thresholds to shrink the WP coefficients dominated by noise.

6.6 Chapter summary

The wavelet transform is a relatively recent mathematical development that has quickly found a place in many engineering applications. The mathematical basis of wavelet theory is presented in this Chapter, as well as several implementation issues. Additionally, this Chapter also discusses heuristic SURE soft threshold shrinkage algorithms for THz noise removal. This Thesis carries out wavelet transform analysis using predominantly Mallat's pyramid algorithm. Wavelet packet transforms of THz measurements discussed in Section 6.4 are especially favoured since wavelet packet bases further improve the flexibility of space and frequency localisation, compared to the critically subsampled DWTs. The performance of WPTs for THz signal recognition will be experimentally illustrated in Chapter 9 (Sec. 9.3.1 and Sec. 9.4). In addition, two dimensional DWTs described in this Chapter are desirable for THz image processing and reconstruction due to its good time-frequency localisation feature, which are experimentally illustrated in Chapters 11-13. The next Chapter investigates a series of different feature extraction methods to achieve the transformed signal characteristics of THz measurements. It is the second phase of the pattern identification system to process acquired THz measurements and to reveal important differences for discrimination between patterns associated with different classes.

Chapter 7

Feature Extraction and Selection

One of the tasks of pattern recognition is to convert patterns to features, where these features are a description of the collected data in a compact form. Ideally, these features only contain relevant information, which then play a crucial role in determining the division of properties concerning each class. Mathematical models of feature extraction lead to a dimensionality reduction, resulting in lower dimensional representation of the information. Following feature extraction, feature selection has an important influence on classification accuracy, necessary time for classification, the number of examples for learning and the cost of performing classification.

Generally speaking, for different pattern recognition problems, there are various solution strategies. In this Chapter, we will discuss five mathematical models for extracting features: Fourier transforms; AR modelling over wavelet decomposition levels; a subspace identification algorithm; a wavelet packet identification procedure; and optimized Mertz apodization functions for system identification. These feature extraction methods have been successfully applied to identify two- and multiclass biomedical samples from their THz measurements, which will be discussed in detail in Chapter 9.

7.1 Role of feature selection and extraction

A pattern recognition system includes feature extraction and selection. The collected data are subject to the feature extraction and selection processing. The aim is to determine the feature sets as input to the subsequent classifier. The pattern processing in the Thesis is mainly about signals collected from THz imaging setups, which are time series data at each pixel. A new question is to explore the identification decision made with respect to these signal vectors, as being conducted in a 3D volume, instead of conventional pattern recognition from 2D or 1D data.

Both feature selection and feature extraction can be applied to accomplish dimensionality reduction for classification. Feature selection omits redundant information and irrelevant features from the available measurements. All the ignored features do not contribute to class separability. Feature extraction, on the other side, considers all the information content and maps the useful information content into a lower dimensional feature space. In this Thesis, we use feature extraction methods to map useful information and then apply feature selection to trace only significant input feature sets to classifiers, which allows us to accomplish classification experiments on THz measurements to desirable accuracy levels.

It is important to achieve dimensionality reduction for effective classification tasks. A finite set of inputs are then applied for the design of a recognition system. A series of mathematical methods have been investigated for optimal classification performance, and is summed up in Table 7.1 from Meyer-Base (2003).

7.2 Feature extraction methods

There are three considerations in the extraction and selection of features: (i) feature evaluation criterion, (ii) dimensionality of the feature space, and (iii) optimisation procedure.

A good feature should satisfy two criteria. Firstly, a feature should remain unchanged when variations take place within a class. The second is the ability of the feature to reveal important differences of the discrimination between patterns associated with different classes. The result is that the good features allow a decision of the target patterns with use of enough pertinent information.

Table 7.1. Standard approaches in pattern recognition.

Measured features	Transformation of features	Structural features
Amplitude	Polynomials	Peaks
Bias	Harmonic analysis	Derivatives
Duration	Fourier tansforms	Lines
Phase	Wavelet transforms	Edges
Energy	Haar transform	LPC coefficients
Moments	Karhunen-Loève Transform	Parametric models
Singular value		
Karhunen-Loève		
Eigenvalues		

Feature selection	Classifiers	Clustering methods
Discriminant analysis	Euclidian distance	Isodata algorithm
Chernoff bound	Mahalanobis distance	Fisher's linear discriminant analysis
Bhattacharya divergence	Linear discriminant functions	Parsing
Exhaustive search	Bayesian linear classifier	
Dynamic programming	Maximum likelihood	
	Production rules	
	Support vector machines	
	Density functions	
	Parzen estimator	
	k-NN algorithm	
	Histogram	

There are two known categories in the literature for feature extraction, which are used in THz pattern recognition: (i) nontransformed structural characteristics: moments, power, amplitude information, energy, etc. (ii) transformed structural characteristics: frequency and amplitude spectra, subspace transformation methods, etc.

The so-called curse of dimensionality (Duda and Hart 1973) is one of the most frequent problems when applying any classification technique to pattern recognition. Methods that are adequate for a low-dimensional feature space might be completely impracticable for a high-dimensional space (number of features > 50). This shows that techniques for dimensionality reduction in the feature space have to be developed in order to solve practical problems.

We will tackle the problem of feature extraction applied to THz biomedical signal analysis. Our goal will be for a given collection of signals to generate the features that will be the input to a classifier. These signals will have to be assigned labels to one of the possible classes.

For the current Chapter, we will discuss the transformed signal characteristics. The relevant feature extraction and selection methods will also be discribed.

7.3 Fourier transform for signal analysis

The Fourier transform has been a standard analytical tool for engineers and scientists over the past two centuries. The Fourier transform defines a relationship between a signal in the time domain and its representation in the frequency domain. In this specific case, both domains are continuous and unbounded. Being a unitary transform, no information is created or lost in the process, so the original signal can be recovered from its Fourier transform. The definitions for the Fourier transform and its inverse are

$$F(\omega) = \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} \mathrm{d}t \tag{7.1}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} \mathrm{d}t \tag{7.2}$$

where the pair of functions $\{f(t), F(\omega)\}$ is known as a Fourier transform pair. The Fourier transform can be viewed as an expansion in the orthonormal basis $\{e^{-j\omega t} | \omega \in R\}$. These basis functions are the eigenfunctions for linear, time-invariant operators. The Fourier transform is an ideal mathematical tool for the study of such systems. The eigenfunction expansion on the signals of interest forms the theoretical foundations for frequency domain analysis that is very important in the mathematical sciences. A frequency orientation component algorithm is introduced later in Section 9.5.2 to achieve THz signal identification.

However, Fourier basis functions are pure sinusoids, and they have infinite support in time. This fact renders Fourier techniques unsuitable for describing signals containing discontinuities and sharp spikes. In addition, in the inverse transform expression of Eq. (7.2), the convolution kernel, $e^{j\omega t}$, has unit magnitude, indicating equal contributions to the inverse transform from every frequency in the transform. The result leads to a lack of compactness for non-stationary signals, which is a major disadvantage in applications such as signal compression. This is the reason that in this Thesis, both Fourier transforms and wavelet transforms are applied for the analysis of THz pulsed signals, and the latter forms the majority of the work in the PhD project. Wavelet transforms have been described in the previous Chapter.

The invention of the Fast Fourier Transform (FFT) has proved to be a significant milestone in the development of Fourier analysis. Along with rapid advances in digital computers, the FFT provides a practically feasible method to compute the Fourier transform. Subsequently, Fourier analysis remains, to this day, an important technique

for scientists and engineers. In this section, Fourier techniques are used for feature extraction methods, allowing a acceptable classification performance in THz measurements. Further applications, i.e. computed tomographic reconstruction algorithms, will be introduced later in Chapter 10.

7.4 AR and ARMA parametisation of wavelet coefficients

In this Section, a novel technique involving the use of Auto Regressive (AR) and Auto Regressive Moving Average (ARMA) models on the wavelet transforms of T-ray pulse responses is used as feature extraction method for classification purposes.

7.4.1 AR model parameter estimation

Since the AR model is built on wavelet subbands, in the subsequent wavelet decomposition, we denote the new approximation and detail coefficients with $\bar{\mathbf{c}}$ and $\bar{\mathbf{d}}$, respectively.

At every level k of the DWT decomposition, we perform firstly auto-regressive (AR) and then moving average (MA) modelling of the coefficients. We use the forward linear prediction (Therrien and Oppenheim 1992, Proakis and Manolakis 1996) at the j^{th} wavelet transform depth given by the linear difference equation in the time domain:

$$\hat{\bar{c}}_j^{\text{AR}}[n] = -a_j[1]\bar{c}_j[n-1] - a_j[2]\bar{c}_j[n-2] - ... - a_j[P]\bar{c}_j[n-P] + w_j[n] \tag{7.3}$$

where $\hat{\bar{c}}_k^{\text{AR}}[n]$ represents the current prediction of wavelet approximation coefficients through AR modelling, and P represents the prediction order of the model. The modelling residual $w_k[n]$ is assumed to be a white Gaussian process. The least-squares form of the Yule-Walker equation is used to estimate the corresponding AR model parameters (Therrien and Oppenheim 1992, Jain and Deshpande 2004):

$$\mathbf{R}_{\bar{\mathbf{c}}_j}\mathbf{a}_j = [S_j \quad \mathbf{0}]^{\text{T}} \tag{7.4}$$

where, $\mathbf{R}_{\bar{\mathbf{c}}_j}$ is the covarianace matrix of the data, $\mathbf{a}_j = [a_j[0], ..., a_j[P]]^{\text{T}}$ is the vector of the AR model coefficients and S_j is the minimum sum of squared errors. The data vector $\bar{\mathbf{c}}_j$ of the autocorrelation method is determined by the selection of the end points as $n_I = 0$, $n_F = N_s + P - 1$, and N_s is selected to be the length of the detailed wavelet coefficient at the j^{th} wavelet transform level, with $N_s = n = L_x/(2 \times j)$, where L_x

labels the length of data vector **x**. The prediction error variance (PEV), which is used to validate our AR model, is calculated as

$$\sigma^2_{\epsilon j,P} = \frac{1}{n_F - n_I + 1} S_j. \tag{7.5}$$

It should be noted that the prediction error variance is the theoretic criterion for any AR/ARMA model order selection task. It measures the precision of a model's predictions. The measurement errors are reduced by the model fitting process if the PEV is less than 1, while any errors in the data measurements are multiplied if the PEV is greater than 1. It is desirable for PEV to be close to zero, in which case, the predictive power of the model would be more accurate. This is a crucial first step to estimate the model parameters (Liang *et al.* 1993).

7.4.2 ARMA model parameter estimation

An ARMA model is a combination of AR and moving average (MA) models. The MA model at the j^{th} wavelet decomposition level, based on approximation coefficients $\bar{\mathbf{c}}_j$ is developed from the following difference equation:

$$\hat{\bar{c}}^{\text{MA}}_j[n] = \bar{c}_j[0]w_j[n] + b_j[1]w_j[n-1] + ... + b_j[Q]w_j[n-Q]. \tag{7.6}$$

The MA prediction output $\hat{\bar{c}}^{\text{MA}}_j[n]$ is seen as a sum of weighted average of the past Q_j input samples of a white noise process $w_j(n)$. The full ARMA model can be obtained by combining both AR and MA models, as below:

$$\sum_{i=0}^{P_j} a_j[i]\bar{c}_j[n-i] = \sum_{k=0}^{Q_j} b_j[k]w_k[n-k] \tag{7.7}$$

where, $a_j[0]$ is normalised to 1.

The MA parameters are estimated by Durbin's method, which is an improvement of the basic Prony method that assumes that $\hat{\bar{\mathbf{c}}}_j[n] = \bar{\mathbf{c}}_j[n]$ for $n = 0, 1, ...N_s - 1$, setting $N_s = P_j + Q_j + 1$ with $P_j = Q_j$. The idea behind Durbin's method is to turn a MA modelling problem into a set of two normalised and significantly over-parametrised AR modelling problems denoted by A and B and consequently solve them (i.e. adopting the condition $N_s > P_j + Q_j + 1$). The purpose of this over-parametrisation process is to facilitate the solution. In our work, the order of AR model (P_j) is selected to be five times the MA model order Q_j ($P_j = 5Q_j$) as this has been found to provide a good fit

(Therrien and Oppenheim 1992). Then the Yule-Walker matrix equation of the ARMA model is partitioned as:

$$\left[\mathbf{R}_{\bar{\mathbf{c}}_{j_B}} \quad \mathbf{R}_{\bar{\mathbf{c}}_{j_A}}\right]^{\mathrm{T}} \mathbf{a}_j = \left[\mathbf{b}_j \quad \mathbf{e}_{j_A}\right]^{\mathrm{T}} \tag{7.8}$$

where $\mathbf{b}_j = [b_j[0]...b_j[Q]]$. Here $\mathbf{e}_{j_A}$ is the prediction error that is defined as $\mathbf{e}_{j_A} = \bar{\mathbf{c}}_j[n] - \hat{\bar{\mathbf{c}}}_j[n]$, with $n = 0, 1, ...N_s - 1$. The related AR coefficient vector $\mathbf{a}_j$ is determined from minimising $S_{j_A} = |\mathbf{e}_{j_A}|^2$ and observing that $\mathbf{R}_{\bar{\mathbf{c}}_{j_A}} \mathbf{a}_j = \mathbf{e}_{j_A}$ is equivalent to the Yule-Walker equations for the AR model in Eq. (7.4). The difference between the two formulations lies in the calculation of MA parameters $\mathbf{b}_j$. Instead of forcing $\mathbf{b}_j$ to match the left side of the equation $\mathbf{R}_{\bar{\mathbf{c}}_{j_B}} \mathbf{a}_j = \mathbf{b}_j$ in Eq. (7.8), Durbin's method considers the existence of an error term of this equation. A higher order AR model is used as a substitute for an MA model. It is worth noting that (i) the coefficient vector $\mathbf{a}_j$ is obtained from the solution of the Yule-Walker equation normalised by the gain—first item of AR coefficients, before substituting S_j with unity in Eq. (7.4); (ii) we use the coefficient vector $\mathbf{a}_j$, derived from setting $P_j = 5Q_j$, to construct the data matrix that provides the new AR modelling coefficients with the MA model order Q_j. Durbin's method also guarantees the stability of the ARMA model eliminating windowing effects.

7.4.3 Feature extraction via AR models over wavelet decomposition

The objective of feature extraction is to isolate the relevant features mentioned earlier from the T-ray signals to improve classification performance. Feature matrix calculations, assuming AR, MA and ARMA models of different order, are performed. The averages of the modelling coefficients are computed over the three decomposition levels of the wavelet transform employed on each data set. The model coefficient averages are then joined to produce feature vectors with a dimension equal to the number of subbands in the wavelet decomposition. The feature vectors obtained from two different AR orders, and MA orders are combined, respectively, to form the final AR and MA feature matrices. The ARMA feature matrix is obtained by combining two different orders of AR and MA vectors together. The extracted AR and MA feature vectors are calculated at each decomposition level j by

$$\langle \mathbf{a}_j^{P^*} \rangle = \frac{1}{P^*} \sum_{i^*=1}^{P^*} a_j[i^*], \text{where } 2 \leq P^* \leq 7 \tag{7.9}$$

$$\langle \mathbf{a}_j^{P} \rangle = \frac{1}{P} \sum_{i=1}^{P} a_j[i], \text{where } 3 \leq P \leq 8 \tag{7.10}$$

$$\langle \mathbf{b}_j^{Q^*} \rangle = \frac{1}{Q^*} \sum_{k^*=1}^{Q^*} b_j[k^*], \text{where } 2 \leq Q^* \leq 7 \tag{7.11}$$

$$\langle \mathbf{b}_j^{Q} \rangle = \frac{1}{Q} \sum_{k=1}^{Q} b_j[k], \text{where } 3 \leq Q \leq 8 \tag{7.12}$$

where $\langle \mathbf{a}_j^P \rangle$, $\langle \mathbf{a}_j^{P^*} \rangle$, and $\langle \mathbf{b}_j^Q \rangle$, $\langle \mathbf{b}_j^{Q^*} \rangle$ denote the averaged value (DC value) of AR and MA model coefficients at different model orders. The new feature matrix using the AR modelling procedure is given from

$$\mathbf{DC}_j^{\text{AR}} = [\langle \mathbf{a}_j^{P^*} \rangle | \langle \mathbf{a}_j^P \rangle], \text{where } P^* \leq P. \tag{7.13}$$

A similar procedure is adopted for the MA modeled data sets

$$\mathbf{DC}_j^{\text{MA}} = [\langle \mathbf{b}_j^{Q^*} \rangle | \langle \mathbf{b}_j^Q \rangle], \text{where } Q^* \leq Q. \tag{7.14}$$

Finally, we obtain new feature matrices combining the coefficients from both the AR and the MA modelling procedures

$$\mathbf{DC}_j^{\text{ARMA}} = [\langle \mathbf{a}_j^{P^*} \rangle | \langle \mathbf{b}_j^Q \rangle]. \tag{7.15}$$

This combination aims at improving the generalization ability of the classification process. Additionally, since the feature matrices are the combination of the averaged coefficients matrix related to AR, MA, and ARMA modelling of wavelet approximation coefficients at three decomposition levels, we relabel the wavelet decomposition level j to a combination of all wave decomposition levels $\mathbf{j}$. Therefore we obtain the new feature matrices: $\mathbf{DC}_\mathbf{j}^{\text{AR}}$, $\mathbf{DC}_\mathbf{j}^{\text{AR}}$ and $\mathbf{DC}_\mathbf{j}^{\text{AR}}$.

The complete procedure for calculating $\mathbf{DC}_\mathbf{j}^{\text{ARMA}}$ is depicted in Fig. 7.1. The implementation used in this Thesis is motivated by a desire to obtain fixed length feature vectors to facilitate comparison. Since various AR/ARMA model orders are used, this naturally leads to a varying number of model coefficients, and hence affects the dimensionality of the extracted feature vectors. Averaging the AR/ARMA coefficients aims to transform the varying dimension of AR/ARMA model coefficients to a fixed number of dimensions. Another parameter that can affect the feature matrix length is the number of decomposition levels in wavelet transforms, since varying $\mathbf{j}$ leads to a different number of subbands.

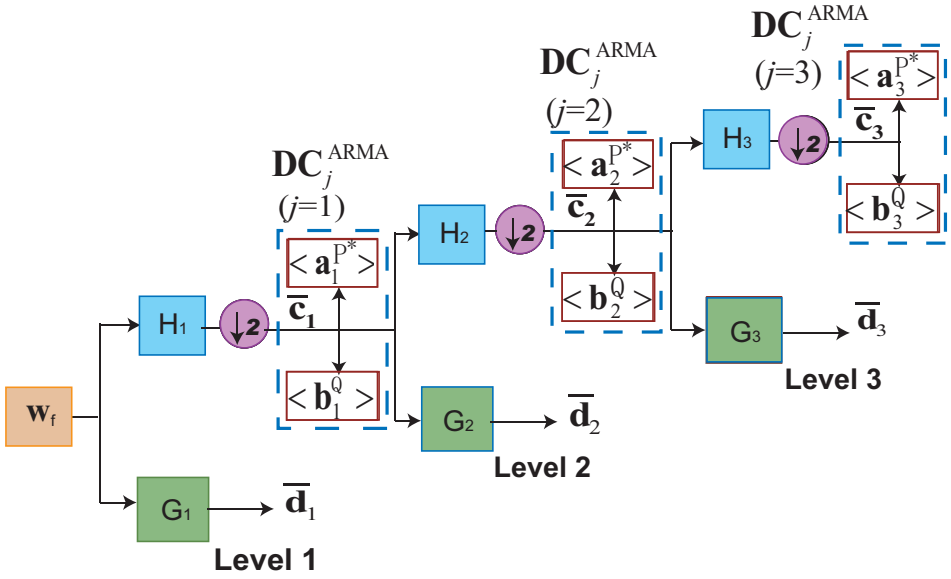


Figure 7.1. Illustration of the DC^{ARMA} feature matrix calculated over three levels of wavelet decomposition. In this modelling, H and G denote the low- and high-pass filters, respectively, $\mathbf{w}_f$ is the de-noised T-ray input. The arrow indicates the diadic downsampling process. Similar illustration related to DC^{AR} and DC^{MA} feature matrix are assumed.

7.5 System identification for feature extraction

System identification refers to the problem of estimating a system that best describes the measured data. System identification starts from experimental data, measurements, and observations. The data are assumed to consist of two sets: the output of the unknown system, which is excited by the input signal. We consider the system whose output depends on the input signal and noise.

Prediction error minimization and subspace identification are two popular approaches to evaluate an identification system. Predictor error methods, such as the ARMA model, represented above, are based on local optimization algorithms, while subspace methods are based on geometric operations on subspaces spanned by matrices obtained from measurements without involving explicit optimization criteria (Van Overschee and De Moor 1994). Hence, in general, subspace methods are fast and numerically stable, however, less accurate than prediction error minimization methods.

In addition to subspace model for system identification, this Chapter also explores two more common models for the identification task. They are wavelet packet identification models and optimized Mertz apodization functions.

For analysing the THz data, the THz pulse detected with no sample in place is considered to be the input, and the THz pulse detected after transmission through the sample is taken as the system output. Via the calculation of system parameters, the time domain pulse is reduced to a relative small number of features. It turns out to be a classifier that can be trained and an unknown sample that can be identified.

7.5.1 A subspace system identification algorithm

The subspace identification approach to a state-space model adopts a time-invariant process, which implies an unchangeable underlying physical system. The model order, which is the only parameter that needs to be chosen, can be conveniently estimated from a singular value plot. Efficient linear algebra tools such as singular value decomposition are applied. The subspace identification algorithm, based on the work carried out by Galvão *et al.* (2005) is applied for THz system identification in this Thesis. The system identification algorithm consists of two modelling subsystems. One is called deterministic modelling of the identification system, where the output of the system only depends on the input signal. The other is called stochastic modelling of the system, where the output depends on a noise signal. Since we consider the system with output depending on the input signal and a noise signal, we derive the identification system based on the combination of deterministic and stochastic models. They are reviewed as follows.

The deterministic subsystem

Discrete-time state-space models are arranged with only one delay involved. This leads to the introduction of extra variables, known as state variables. The state variables can be reconstructed from input-output data. The number of state variables needed in the representation is equal to the order of the system. A single-input single-output state-space model can be written as

$$\begin{aligned} \mathbf{x}[k+1] &= \mathbf{A}\mathbf{x}[k] + \mathbf{B}u[k] \\ y[k] &= \mathbf{C}\mathbf{x}[k] + Du[k] \end{aligned} \tag{7.16}$$

where $\mathbf{x}[k]$ is an n-dimensional state column vector, $u[k]$ and $y[k]$ are the input and the output, respectively. In addition, this equation is represented by a scalar parameter D and matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ with the dimensions $(n \times n)$, $(n \times 1)$ and $(1 \times n)$, respectively. There is an equivalence between the state-space representation and the input-output representation. The equation is described via taking the z-transform of the state-space model with eliminating state variables

$$G_d(z) = \frac{Y(z)}{U(z)} = \mathbf{C}(z\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + D \tag{7.17}$$

where $\mathbf{I}$ is an identity matrix. The estimation of the $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ matrices and the scalar D is performed via applying the subspace approach. There are three steps: (i) the determination of the model order via a singular value plot, (ii) the estimation of the state sequence from the input-output data via a generalized observability matrix, and (iii) the estimation of system matrices via a linear regression procedure, with use of the input, output, and state sequences.

In order to achieve the possible representations of state matrices from the observed input-output relationship, the decoupled modal realisations are of interest wherein there is no energy exchange between the propagating modes.

In the subspace identification procedure, the input data are arranged in Hankel matrices as

$$\mathbf{U}_{0|2i-i} = \begin{bmatrix} u[0] & u[1] & u[2] & \cdots & u[j-1] \\ u[1] & u[2] & u[3] & \cdots & u[j] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \hline u[i] & u[i+1] & u[i+2] & \cdots & u[i+j-1] \\ u[i+1] & u[i+2] & u[i+3] & \cdots & u[i+j] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u[2i-1] & u[2i] & u[2i+1] & \cdots & u[2i+j-2] \end{bmatrix} \tag{7.18}$$

where the notation $\mathbf{U}_{0|2i-1}$ implies that the first row of the matrix begins with $u[0]$ and the last row begins with $u[2i-1]$. There is an assumption that $j \to \infty$ in the Hankel matrix for identification. The past input is denoted with $\mathbf{U}_p$, and the future input is denoted with $\mathbf{U}_f$, both of which are divided by a reference point $k = i$. The division i is artificially imposed.

In a similar fashion, the output data are arranged in a Hankel matrix $\mathbf{Y}$, with $\mathbf{Y}_p$ and $\mathbf{Y}_f$. Using this matrix notation, the system equations are as follows:

$$\mathbf{Y}_p = \mathbf{\Gamma}_i \mathbf{X}_p + \mathbf{\Omega}_i \mathbf{U}_p \tag{7.19}$$

$$\mathbf{Y}_f = \mathbf{\Gamma}_i \mathbf{X}_f + \mathbf{\Omega}_i \mathbf{U}_f \tag{7.20}$$

$$\mathbf{X}_f = \mathbf{A}^i \mathbf{X}_p + \mathbf{\Delta}_i \mathbf{U}_p \tag{7.21}$$

where $\mathbf{X}_p = [\mathbf{x}[0]\ \mathbf{x}[1] \cdots \mathbf{x}[i-1]]$, $\mathbf{X}_f = [\mathbf{x}[i]\ \mathbf{x}[i+1] \cdots x[i+j-1]]$, and $\mathbf{\Gamma}_i$ is an extended observability matrix built as

$$\mathbf{\Gamma}_i = \begin{bmatrix} \mathbf{C} & \mathbf{CA} & \mathbf{CA}^2 & \cdots & \mathbf{CA}^{i-1} \end{bmatrix}^{\mathrm{T}} \tag{7.22}$$

(where T signifies transpose). This relates the state sequence with the input-output data, and $\mathbf{\Omega}_i$ is the impulse response matrix given by

$$\mathbf{\Omega}_i = \begin{bmatrix} \mathbf{D} & 0 & 0 & 0 & 0 \\ \mathbf{CB} & \mathbf{D} & 0 & \cdots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{CA}^{i-2}\mathbf{B} & \mathbf{CA}^{i-3}\mathbf{B} & \mathbf{CA}^{i-4}\mathbf{B} & \cdots & \mathbf{D} \end{bmatrix} \tag{7.23}$$

and $\mathbf{\Delta}_i$ is an extended controllability matrix given by $\mathbf{\Delta}_i = [\mathbf{A}^{i-1}\mathbf{B}\ \mathbf{A}^{i-2}\mathbf{B}\ \cdots\ \mathbf{AB}\ \mathbf{B}]$. As the system is assumed to be observable and controllable, the unknown matrices $\mathbf{\Gamma}_i$ and $\mathbf{\Delta}_i$ have full rank. The state sequence can be obtained from

$$\mathbf{O}_i = \mathbf{\Gamma}_i \mathbf{X}_f \tag{7.24}$$

where $\mathbf{O}_i$ is the oblique projection (see the Appendix A) of $\mathbf{Y}_f$ onto $\mathbf{W}_p$ along the $\mathbf{U}_f$ direction, denoted

$$\mathbf{O}_i = \mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p \tag{7.25}$$

where matrix $\mathbf{W}_p$ is built by stacking the past input and output Hankel matrices as $\mathbf{W}_p = [\mathbf{U}_p^{\mathrm{T}}, \mathbf{Y}_p^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{O}_i$ are the components of the system output, which result from excitations that took place before the reference instant i. After $\mathbf{O}_i$ has been obtained, we need to factorize it into two terms $\mathbf{\Gamma}_i$ and $\mathbf{X}_i$.

The singular value decomposition (SVD) of $\mathbf{O}_i$ can be implemented to achieve a numerically stable factorization. Additionally, it also provides a way of determining the order of a system, which is the rank of $\mathbf{O}_i$. The SVD of $\mathbf{O}_i$ is written as

$$\mathbf{O}_i = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} = \begin{pmatrix} \mathbf{U}1 & \mathbf{U}2 \end{pmatrix} \begin{pmatrix} \mathbf{S}_1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_2^{\mathrm{T}} \end{pmatrix} = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^{\mathrm{T}} \tag{7.26}$$

where $\mathbf{S}_1$ is a diagonal matrix containing singular values different from zero and $\mathbf{U}_1$, $\mathbf{V}_1$ are unitary matrices. From Eq. (7.26), it can be observed that any choice of $\mathbf{\Gamma}_i$ and $\mathbf{X}_f$ such that $\mathbf{\Gamma}_i\mathbf{X}_f = \mathbf{U}_1\mathbf{S}_1\mathbf{V}_1^{\mathrm{T}}$ is a solution to the factorization. The diagonal elements in S_1 are related to the variance along the principal directions of O_i.

To balance the variance between $\mathbf{\Gamma}_i$ and $\mathbf{X}_f$, we split the $\mathbf{S}_1$ matrix into two equal terms, and therefore, the balanced solution can be expressed $\mathbf{\Gamma}_i = \mathbf{U}_1\mathbf{S}_1^{1/2}\mathbf{T}$ and $\mathbf{X}_f = \mathbf{T}^{-1}\mathbf{S}_1^{1/2}\mathbf{V}_1^{\mathrm{T}}$. By choosing different matrices $\mathbf{T}$ we achieve different state space realizations of the input-output relation. By estimating the entire state sequence at once, the subspace method avoids the propagation of errors that are associated with the recursive computation of the states from an initial estimate. The geometric analysis for this identification procedure via Hankel matrices can be referred to in Appendix A.

After the state sequence is obtained, it can be used in conjunction with the input and output data to estimate the system matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$. This is carried out by a matrix linear regression procedure based on the following equation:

$$\begin{bmatrix} \mathbf{X}_{i+1} \\ \mathbf{Y}_{i|i} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{X}_i \\ \mathbf{U}_{i|i} \end{bmatrix} \tag{7.27}$$

where $\mathbf{Y}_{i|i} = \{y[i], y[i+1] \cdots y[i+j-1]\}$, $\mathbf{U}_{i|i} = \{u[i], u[i+1] \cdots u[i+j-1]\}$, $\mathbf{X}_i = \mathbf{X}_f$ and $\mathbf{X}_{i+1} = \{x[i+1], x[i+2], \cdots, x[i+j]\}$ is the shifted state sequence, which can be obtained by an oblique projection procedure (see Appendix A) from the equation

$$\mathbf{Y}_f^- = \mathbf{\Gamma}_{i-1}\mathbf{X}_{i+1} + \mathbf{\Omega}_{i-1}\mathbf{U}_f^- \tag{7.28}$$

where the minus superscript implies that the first row of the matrix has been removed.

Once $\mathbf{X}_{i+1}$ has been estimated, Eq. (7.27) can be solved for $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ by a least-squares procedure by post-multiplying both sides of the equation by the pseudo-inverse of $[\mathbf{X}_i^{\mathrm{T}}, \mathbf{U}_{i|i}^{\mathrm{T}}]^{\mathrm{T}}$. In the absence of noise, the system of equations described in Eq. (7.27) is consistent and as a result there are no residuals in the least-squares solution.

Stochastic subsystem

The stochastic state $\mathbf{x}^s[k]$ and output $y^s[k]$ follow from the stochastic subsystem, which describe the influence of the noise sequences $\mathbf{e}_\mathbf{x}[k_1]$ and $e_y[k_2]$ on the stochastic output:

$$\mathbf{x}[k+1]^s = \mathbf{A}\mathbf{x}^s[k] + \mathbf{e}_\mathbf{x}[k] \tag{7.29}$$

$$y^s[k] = \mathbf{C}\mathbf{x}^s[k] + e_y[k]. \tag{7.30}$$

Here, $\mathbf{e}_x[k]$ and $e_y[k]$ are assumed to be white, have zero mean and covariance matrices satisfying the following equation

$$\mathbb{E}\left(\begin{bmatrix} \mathbf{e}_\mathbf{X}[k_1] \\ e_y[k_2] \end{bmatrix} \begin{bmatrix} \mathbf{e}_\mathbf{X}^\mathrm{T}[k_1] & e_y[k_2] \end{bmatrix} \right) = \begin{bmatrix} \mathbf{Q} & \delta(k_1,k_2)\mathbf{S} \\ \delta(k_1,k_2)\mathbf{S}^\mathrm{T} & R \end{bmatrix} \tag{7.31}$$

where $\delta(k_1,k_2) = 1$ if $k_1 = k_2$ and 0 otherwise.

For the stochastic subsystem, we define

$$\begin{aligned} \mathbf{P} &\underline{\underline{\text{def}}}\ \mathbb{E}\{\mathbf{x}^s[k](\mathbf{x}^s[k])^\mathrm{T}\} \\ \mathbf{G} &\underline{\underline{\text{def}}}\ \mathbb{E}\{\mathbf{x}^s[k](y^s[k])^\mathrm{T}\} \\ \mathbf{\Lambda}_0 &\underline{\underline{\text{def}}}\ \mathbb{E}\{y^s[k](\mathbf{x}^s[k])^\mathrm{T}\}. \end{aligned} \tag{7.32}$$

With Eqs. (7.29), (7.30), (7.31), and through stability of the controllable modes of the system $\{\mathbf{A}, \mathbf{Q}^{1/2}\}$, the following equations are satisfied

$$\begin{aligned} \mathbf{P} &= \mathbf{A}\mathbf{P}\mathbf{A}^\mathrm{T} + \mathbf{Q} \\ \mathbf{G} &= \mathbf{A}\mathbf{P}\mathbf{C}^\mathrm{T} + \mathbf{S} \\ \mathbf{\Lambda}_0 &= \mathbf{C}\mathbf{P}\mathbf{C}^\mathrm{T} + R. \end{aligned} \tag{7.33}$$

This set of equations describe the set of all possible stochastic realisations that have the same second order statistics as a given stochastic sequence $y^s[k]$. More details can be referred to in Faure (1976).

It is easy to derive

$$\mathbf{\Lambda}_i\ \underline{\underline{\text{def}}}\ \mathbb{E}\{y^s[k+i](y^s[k])^\mathrm{T}\} = \begin{cases} \mathbf{C}\mathbf{A}^{i-1}\mathbf{G} & \text{if } i > 0 \\ \mathbf{\Lambda}_0 & \text{if } i = 0 \\ \mathbf{G}^\mathrm{T}(\mathbf{A}^\mathrm{T})^{-i-1}\mathbf{C}^\mathrm{T} & \text{if } i < 0. \end{cases} \tag{7.34}$$

Associated with the stochastic subsystem, we define the following matrices: (i) the matrix $\mathbf{\Delta}_i^s$

$$\mathbf{\Delta}_i^s \underline{\underline{\text{def}}} \left(\mathbf{A}^{i-1}\mathbf{G} \;\; \mathbf{A}^{i-2}\mathbf{G} \;\; \cdots \;\; \mathbf{A}\mathbf{G} \;\; \mathbf{G} \right), \tag{7.35}$$

(ii) the block Toeplitz covariance matrix $\mathbf{L}_i^s$

$$\mathbf{L}_i^s \underline{\underline{\text{def}}} \begin{bmatrix} \mathbf{\Lambda}_0 & \mathbf{\Lambda}_{-1} & \mathbf{\Lambda}_{-2} & \cdots & \mathbf{\Lambda}_{1-i} \\ \mathbf{\Lambda}_1 & \mathbf{\Lambda}_0 & \mathbf{\Lambda}_{-1} & \cdots & \mathbf{\Lambda}_{2-i} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Lambda}_{i-1} & \mathbf{\Lambda}_{i-2} & \mathbf{\Lambda}_{i-3} & \cdots & \mathbf{\Lambda}_0 \end{bmatrix}, \tag{7.36}$$

and (iii) the block Toeplitz cross covariance matrix $\mathbf{H}_i^s$

$$\mathbf{H}_i^s \underline{\underline{\text{def}}} \begin{bmatrix} \mathbf{\Lambda}_i & \mathbf{\Lambda}_{i-1} & \mathbf{\Lambda}_{i-2} & \cdots & \mathbf{\Lambda}_1 \\ \mathbf{\Lambda}_{i+1} & \mathbf{\Lambda}_i & \mathbf{\Lambda}_{i-1} & \cdots & \mathbf{\Lambda}_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Lambda}_{2i-1} & \mathbf{\Lambda}_{2i-2} & \mathbf{\Lambda}_{2i-3} & \cdots & \mathbf{\Lambda}_i \end{bmatrix} = \mathbf{\Gamma}_i' \mathbf{\Delta}_i^s. \tag{7.37}$$

For the stochastic subsystem, due to stationarity of $y^s[k]$, the following equations hold true

$$\lim_{j\to\infty} \frac{1}{j} \begin{bmatrix} \mathbf{Y}_{0|i-1}^s \\ \mathbf{Y}_{i|2i-1}^s \end{bmatrix} \left((\mathbf{Y}_{0|i-1}^s)^{\mathrm{T}} (\mathbf{Y}_{i|2i-1}^s)^{\mathrm{T}} \right) = \begin{bmatrix} \mathbf{L}_i^s & (\mathbf{H}_i^s)^{\mathrm{T}} \\ \mathbf{H}_i^s & \mathbf{L}_i^s \end{bmatrix}. \tag{7.38}$$

The above derivation refers to the work conducted by Van Overschee and De Moor (1994) and Van Overschee and De Moor (1995).

Deterministic-stochastic models

If the identification is carried out from noisy input-output measurements, the state-space model can be described as the combination of deterministic-stochastic models, which is formulated by adding noise terms to deterministic subsystem

$$\begin{aligned} \mathbf{x}[k+1] &= \mathbf{A}\mathbf{x}[k] + \mathbf{B}u[k] + \mathbf{e}_{\mathbf{x}}[k] \\ y[k] &= \mathbf{C}\mathbf{x}[k] + Du[k] + e_y[k] \end{aligned} \tag{7.39}$$

where column vector $\mathbf{e}_{\mathbf{x}}[k]$ is a state noise, and scalar $e_y[k]$ is an output noise. Here, $\mathbf{e}_{\mathbf{x}}[k]$ and $e_y[k]$ satisfy Eq. 7.31. In this stochastic case, the noise covariance matrices $\mathbf{Q}$, $\mathbf{S}$, R are calculated as a by-product of the identification of the system after the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and scalar D are estimated.

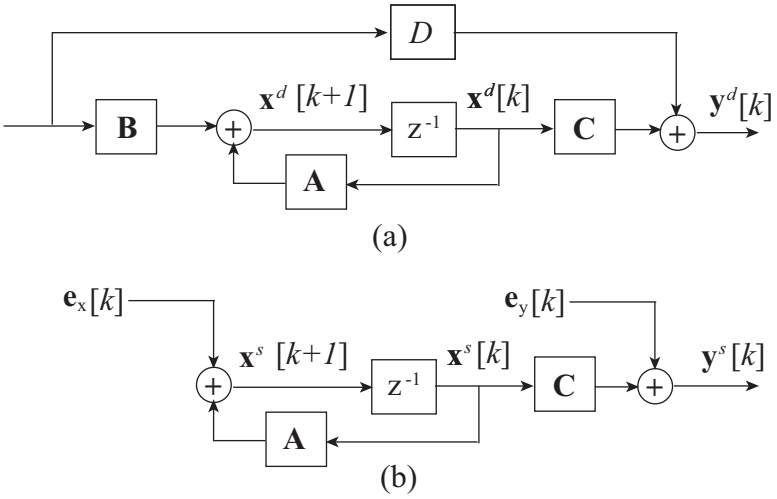


Figure 7.2. Block diagrams of noisy state-space models. (a) Deterministic and (b) stochastic subsystems. Symbol z^{-1} denotes the unit time delay operator (with a slight abuse of notation since the block diagrams depict time-domain signals and not z-transforms). After Galvão *et al.* (2005).

The basic idea of the subspace identification in this stochastic framework is to split the model into two subsystems: deterministic (Fig.7.2(a)) and stochastic. The stochastic subsystem is purely driven by noise (Fig.7.2(b)). The stochastic state $\mathbf{x}^s[k]$ and output $y^s[k]$ follow from the stochastic subsystem, which describe the influence of the noise sequences $\mathbf{e}_\mathbf{x}[k_1]$ and $e_y[k_2]$ on the stochastic output:

$$\mathbf{x}^s[k+1] = \mathbf{A}\mathbf{x}^s[k] + \mathbf{e}_\mathbf{x}[k] \tag{7.40}$$

$$y^s[k] = \mathbf{C}\mathbf{x}^s[k] + e_y[k]. \tag{7.41}$$

Since the system is assumed to be linear, the two subsystems can be combined by obeying the superposition principle. In this case, the system order is only equal to the number of singular values different from zero, achieved via decomposing the oblique projection $\mathbf{O}_i$. The input-output data, in this case, should be acquired in an infinite time. Compared to the value related to system dynamics, the noise associated singular values are quite small. Therefore, statistical criteria can be used to estimate the system order by counting the number of singular values that are significantly larger than zero. In this sense, the identification can be seen as a de-noising process when the noise is placed in a state space.

With superscripts d and s denoting the quantities associated with the deterministic and stochastic subsystems, respectively, the relations between input, output and states can

be expressed as

$$\mathbf{Y}_p = \mathbf{\Gamma}_i \mathbf{X}_p^d + \mathbf{\Omega}_i \mathbf{U}_p + \mathbf{Y}_p^s \tag{7.42}$$

$$\mathbf{Y}_f = \mathbf{\Gamma}_i \mathbf{X}_f^d + \mathbf{\Omega}_i \mathbf{U}_f + \mathbf{Y}_f^s \tag{7.43}$$

$$\mathbf{X}_f^d = \mathbf{A}^i \mathbf{X}_p^d + \mathbf{\Delta}_i \mathbf{U}_p. \tag{7.44}$$

The observability matrix $\mathbf{\Gamma}_i$ and an estimate $\tilde{\mathbf{X}}_f$ of the state sequence can be obtained in a similar manner to that in the deterministic framework, that is: $\mathbf{\Gamma}_i = \mathbf{U}_1 \mathbf{S}_1^{1/2} \mathbf{T}$ and $\tilde{\mathbf{X}}_f = \mathbf{T}^{-1} \mathbf{S}_1^{1/2} \mathbf{V}_1^{\mathrm{T}}$, where $\mathbf{U}_1, \mathbf{S}_1, \mathbf{V}_1$ result from the SVD of $\mathbf{O}_i = \mathbf{Y}_f / \mathbf{U}_f \mathbf{W}_p$ as

$$\mathbf{O}_i = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} = \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{pmatrix} \begin{pmatrix} \mathbf{S}_1 & 0 \\ 0 & \mathbf{S}_2 \end{pmatrix} \begin{pmatrix} \mathbf{V}_1^{\mathrm{T}} \\ \mathbf{V}_2^{\mathrm{T}} \end{pmatrix}. \tag{7.45}$$

Owing to the presence of noise, the singular values in the diagonal matrix $\mathbf{S}$ will all be different from zero. In this case, the separation between $\mathbf{S}_1$ (which contains singular values related to the noise plus the system dynamics) and $\mathbf{S}_2$ (which contains singular values that are related to noise only) needs to be accomplished by inspecting the singular value plot and choosing an appropriate threshold value on the basis of some statistical criterion.

After the state sequence has been estimated, a least squares procedure can be used to estimate the system matrices from the matrix equation

$$\begin{bmatrix} \tilde{\mathbf{X}}_{i+1} \\ \mathbf{Y}_{i|i} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}}_i \\ \mathbf{U}_{i|i} \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{e}}_{\mathbf{x}} \\ \tilde{e}_y \end{bmatrix} \tag{7.46}$$

where $\tilde{\mathbf{e}}_{\mathbf{x}}$ and $\tilde{e}_y$ are residual sequences which can be used to estimate the noise covariance as

$$\begin{bmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^{\mathrm{T}} & R \end{bmatrix} = \mathbb{E} \begin{bmatrix} \tilde{\mathbf{e}}_{\mathbf{x}}^{\mathrm{T}} \\ \tilde{e}_y \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{e}}_{\mathbf{x}}^{\mathrm{T}} & \tilde{e}_y \end{bmatrix}. \tag{7.47}$$

7.5.2 Wavelet-packet identification of a system

The current work exploits the spectral partitioning flexibility of wavelet packets in a subband system identification framework, with link to THz pulses as an illumination source. It is an extension of the work performed by Paiva and Galvão (2006).

Subband model identification

Defining the background and sample responses as the input and output signals, the frequency response of an identified model would be an estimate of the complex insertion loss (CIL). A wavelet packet decomposed tree is adopted to establish the frequency bands at which the subband models will be created. Fig. 6.6 illustrates a wavelet packet decomposition tree in the time domain. Its frequency response is suited to the current wavelet packet scheme (a z-transform of input signals and relative filters). Fig. 7.3 is a particular structure of the wavelet packet decomposition tree, with the application of the proposed scheme. The sub-band model $M_{i,j}(z)$ represents the plant in the frequency band associated to the leaf node (i, j). If (i, j) is a leaf node, then signal $\breve{\mathbf{u}}_{i,j}$ is defined as the output of model $M_{i,j}$ for input $\mathbf{u}$. If (i, j) is not a leaf node, then the coefficients $\mathbf{x}_{i,j}\{\breve{u}_{i,j}\}$ are defined as the reconstruction of the coefficients at the child nodes of (i, j). The sub-band models $M_{i,j}(z)$ are identified from the sample and background responses by following a least-squares procedure as indicated in Fig. 7.4. The least-squares procedure consists of the input signal $\mathbf{u}$ used for identification, model outputs of the plant $\mathbf{y}$ and subband $\breve{\mathbf{u}}_{i,j}$, respectively. Residue $\mathbf{e}_{i,j}$ denotes the wavelet packet coefficients of the difference between $\mathbf{y}$ and $\breve{\mathbf{u}}_{i,j}$, in the frequency band under consideration.

The structure adopted for the subband model is a transfer function of the form

$$M_{i,j}(z) = P_{i,j}(z)Q_{i,j}(z) \tag{7.48}$$

where

$$\begin{aligned} P_{i,j}(z) &= \left(\frac{1}{1-z^{-1}}\right)^{s_{i,j}} \quad s_{i,j} \in \mathbb{Z} \\ Q_{i,j}(z) &= \alpha_{i,j} + \beta_{i,j}z^{-1} \quad \alpha_{i,j},\ \beta_{i,j} \in \mathbb{R} \end{aligned} \tag{7.49}$$

where, $P_{i,j}(z)$ is aimed at roughly approximating the band-limited frequency response of the plant, whereas the FIR term $Q_{i,j}(z)$ provides a fine-tuning for the approximation. In order to improve the match between plant and model, only two taps are used in $Q_{i,j}(z)$, represented in Eq. 7.49. This is reasonable since the system is only represented on a limited frequency band.

A least-squares adjustment for the parameters of $M_{i,j}$ can be carried out by minimizing the following cost function $\mathbf{W}_{i,j} : \mathbb{Z} \times \mathbb{R}^2 \to \mathbb{R}$:

$$\mathbf{W}_{i,j}(s_{i,j}, \alpha_{i,j}, \beta_{i,j}) = \mathbf{e}_{i,j}(\mathbf{e}_{i,j})^{\mathrm{T}} \tag{7.50}$$

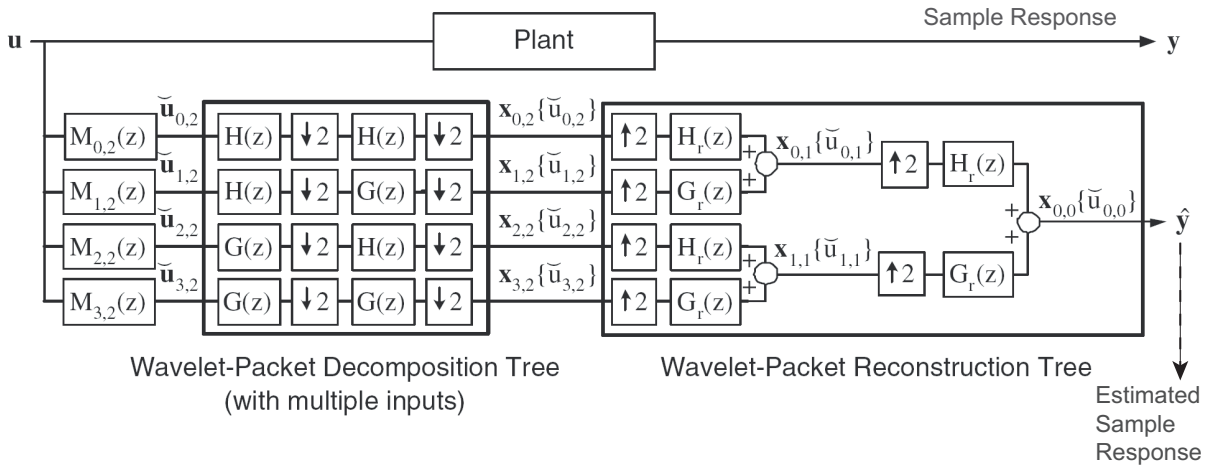


Figure 7.3. Wavelet-packet model structure. In this example, a complete two-level decomposition tree, which defines four frequency sub-bands, is employed. Here, $H(z)$ and $G(z)$ denote low-pass and high-pass decomposition filters, respectively, with reconstruction counterparts represented by $H_r(z)$, $G_r(z)$, which are the relabeled $\tilde{H}$ and $\tilde{G}$. The four sub-band models are represented by the transfer functions $M_{0,2}(z)$, $M_{1,2}(z)$, $M_{2,2}(z)$, $M_{3,2}(z)$. After Paiva and Galvão (2006).

where $\mathbf{e}_{i,j}$ denotes the row vector of residues for the identification data, as in Fig. 7.4. If $s_{i,j}$ is fixed, the optimal real-valued parameters, $\alpha_{i,j}$ and $\beta_{i,j}$, are obtained by imposing

$$\frac{\partial \mathbf{W}_{i,j}}{\partial \alpha_{i,j}} = 0, \quad \frac{\partial \mathbf{W}_{i,j}}{\partial \beta_{i,j}} = 0 \Rightarrow$$

$$\mathbf{e}_{i,j}\left(\frac{\partial \mathbf{e}_{i,j}}{\partial \alpha_{i,j}}\right)^{\mathrm{T}} = 0, \quad \mathbf{e}_{i,j}\left(\frac{\partial \mathbf{e}_{i,j}}{\partial \beta_{i,j}}\right)^{\mathrm{T}} = 0. \tag{7.51}$$

Let $\mathbf{u}_{i,j}^{P}$ be the output of the $P(z)$ term of the subband model for input $\mathbf{u}$ and let $\mathbf{u}_{i,j}^{Pd}$ be the result of delaying $\mathbf{u}_{i,j}^{P}$ by one sample. The model output $\breve{\mathbf{u}}_{i,j}$ can then be written as $\breve{\mathbf{u}}_{i,j} = \alpha_{i,j}\mathbf{u}_{i,j}^{P} + \beta_{i,j}\mathbf{u}_{i,j}^{Pd}$ and thus residue $\mathbf{e}_{i,j}$ becomes

$$\begin{aligned} \mathbf{e}_{i,j} &= \mathbf{x}_{i,j}(\{y\} - \alpha_{i,j}\{u_{i,j}^{P}\} - \beta_{i,j}\{u_{i,j}^{Pd}\}) \\ &= \mathbf{x}_{i,j}\{y\} - \alpha_{i,j}\mathbf{x}_{i,j}u_{i,j}^{P} - \beta_{i,j}\mathbf{x}_{i,j}\{u_{i,j}^{Pd}\}. \end{aligned} \tag{7.52}$$

From Eqs. (7.51) and (7.52), it follows that:

$$(\mathbf{x}_{i,j}\{y\} - \alpha_{i,j}^{*}\mathbf{x}_{i,j}\{u_{i,j}^{P}\} - \beta_{i,j}^{*}\mathbf{x}_{i,j}\{u_{i,j}^{Pd}\})(\mathbf{x}_{i,j}\{u_{i,j}^{P}\})^{\mathrm{T}} = 0 \tag{7.53}$$

$$(\mathbf{x}_{i,j}\{y\} - \alpha_{i,j}^{*}\mathbf{x}_{i,j}\{u_{i,j}^{P}\} - \beta_{i,j}^{*}\mathbf{x}_{i,j}\{u_{i,j}^{Pd}\})(\mathbf{x}_{i,j}\{u_{i,j}^{Pd}\})^{\mathrm{T}} = 0 \tag{7.54}$$

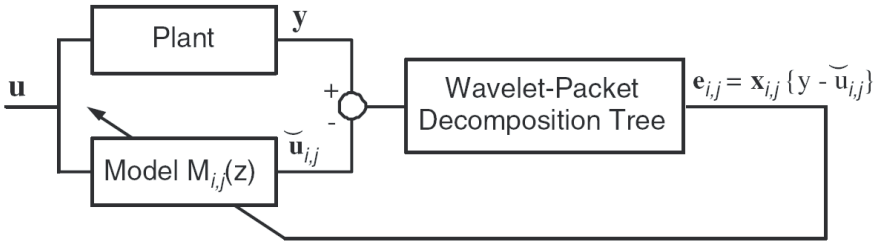


Figure 7.4. Model identification of a sample response for a given frequency sub-band. It illustrates a least-squares procedure that can be applied to achieve the identification of the sub-band models $M_{i,j}(z)$ from the sample and background responses. After Paiva and Galvão (2006).

which leads to

$$\begin{bmatrix} \alpha^* \\ \beta^* \end{bmatrix} = \mathcal{M}^{-1} \begin{bmatrix} \mathbf{x}_{i,j}\{y\}(\mathbf{x}_{i,j}\{u^P_{i,j}\})^{\mathrm{T}} \\ \mathbf{x}_{i,j}\{y\}(\mathbf{x}_{i,j}\{u^{Pd}_{i,j}\})^{\mathrm{T}} \end{bmatrix}, \tag{7.55}$$

where

$$\mathcal{M} = \begin{bmatrix} \mathbf{x}_{i,j}\{u^P_{i,j}\}(\mathbf{x}_{i,j}\{u^P_{i,j}\})^{\mathrm{T}} & \mathbf{x}_{i,j}\{u^{Pd}_{i,j}\}(\mathbf{x}_{i,j}\{u^P_{i,j}\})^{\mathrm{T}} \\ \mathbf{x}_{i,j}\{u^P_{i,j}\}(\mathbf{x}_{i,j}\{u^{Pd}_{i,j}\})^{\mathrm{T}} & \mathbf{x}_{i,j}\{u^{Pd}_{i,j}\}(\mathbf{x}_{i,j}\{u^{Pd}_{i,j}\})^{\mathrm{T}} \end{bmatrix} \tag{7.56}$$

provided that the indicated inverse matrix $\mathcal{M}^{-1}$ exists.

To find the optimal value of $s_{i,j}$, the following search algorithm in $\mathbb{Z}$ is used: the value of $s_{i,j}$ is varied in a specified range. For each value of $s_{i,j}$, the optimal values $\alpha^*_{i,j}$ and $\beta^*_{i,j}$ are calculated using Eq. (7.55), and the corresponding value of $\mathbf{W}_{i,j}$ is noted. The value $s_{i,j}$ for which $\mathbf{W}_{i,j}$ is minimum is then adopted, as well as the corresponding values of $\alpha^*_{i,j}$ and $\beta^*_{i,j}$.

Cost relations along the tree

As for a leaf node (i, j), signal $\breve{\mathbf{u}}_{i,j}$ is the output of subband model $M_{i,j}$ for input $\mathbf{u}$. Since there is no defined subband model $M_{i,j}$ for a non-leaf node, the coefficients $\mathbf{x}_{i,j}\{\breve{u}_{i,j}\}$ are defined as the reconstruction of the coefficients at the child nodes $(2i, j+1)$ and $(2i+1, j+1)$ (see Fig. 7.3 for an example), that is

$$\mathbf{x}_{i,j}\{\breve{u}_{i,j}\} = \mathbf{x}_{2i,j+1}\{\breve{u}_{2i,j+1}\}\mathcal{H}_r + \mathbf{x}_{2i+1,j+1}\{\breve{u}_{2i+1,j+1}\}\mathcal{G}_r \tag{7.57}$$

where $\mathcal{H}_r$ and $\mathcal{G}_r$ are the time-domain reconstruction matrices corresponding to the orthonormal filters H_r and G_r (Vetterli and Kovacevic 1995). Therefore, we define the

matrix $\mathcal{V}_r$ as

$$\mathcal{V}_r = \begin{bmatrix} \mathcal{H}_r \\ \mathcal{G}_r \end{bmatrix} \tag{7.58}$$

where $\mathcal{V}_r(\mathcal{V}_r)^{\mathrm{T}} = (\mathcal{V}_r)^{\mathrm{T}}\mathcal{V}_r = \mathcal{I}$, since it is orthogonal.

The reconstruction can be also applied to signal $\mathbf{y}$

$$\mathbf{x}_{i,j}\{y\} = \mathbf{x}_{2i,j+1}\{y\}\mathcal{H}_r + \mathbf{x}_{2i+1,j+1}\{y\}\mathcal{G}_r. \tag{7.59}$$

In Eq. (7.59), the coefficients $\breve{\mathbf{u}}_{i,j}$, $\breve{\mathbf{u}}_{2i,j+1}$ and $\breve{\mathbf{u}}_{2i+1,j+1}$ indicate three different signals, the last two of which are obtained independently. The modelling process, where $(2i, j+1)$ and $(2i+1, j+1)$ are leaf nodes, will be kept as a best wavelet packet basis, which will be further discussed in the later subsection.

If $(2i, j+1)$ and $(2i+1, j+1)$ are not leaf nodes, then the coefficients of these signals are obtained through Eq. (7.57). Each of these coefficients is the function of its own child nodes and calculated independently. The residue $\mathbf{e}_{i,j}$ for a non-leaf node (i, j) is defined with the same expression used for the leaf nodes. According to Eqs. (7.57), (7.58) and (7.59), we have

$$\begin{aligned} \mathbf{e}_{i,j} &= \mathbf{x}_{i,j}\{y\} - \mathbf{x}_{i,j}\{\breve{u}_{i,j}\} \\ &= \mathbf{e}_{2i,j+1}\mathcal{H}_r + \mathbf{e}_{2i+1,j+1}\mathcal{G}_r \\ &= [\mathbf{e}_{2i,j+1}|\mathbf{e}_{2i+1,j+1}]\mathcal{V}_r. \end{aligned} \tag{7.60}$$

Therefore, the cost

$$\begin{aligned} \mathbf{W}_{i,j} &= \mathbf{e}_{i,j}(\mathbf{e}_{i,j})^{\mathrm{T}} = [\mathbf{e}_{2i,j+1}|\mathbf{e}_{2i+1,j+1}]\mathcal{V}_r\mathcal{V}_r^{\mathrm{T}} \\ &\quad \times [\mathbf{e}_{2i,j+1}|\mathbf{e}_{2i+1,j+1}]^{\mathrm{T}} \\ &= [\mathbf{e}_{2i,j+1}|\mathbf{e}_{2i+1,j+1}][\mathbf{e}_{2i,j+1}|\mathbf{e}_{2i+1,j+1}]^{\mathrm{T}} \\ &= \mathbf{e}_{2i,j+1}(\mathbf{e}_{2i,j+1})^{\mathrm{T}} + \mathbf{e}_{2i+1,j+1}(\mathbf{e}_{2i+1,j+1})^{\mathrm{T}} \\ &= \mathbf{W}_{2i,j+1} + \mathbf{W}_{2i+1,j+1}. \end{aligned} \tag{7.61}$$

That is, cost $\mathbf{W}_{i,j}$ at a non-leaf node (i, j) is equal to the sum of the costs at its child nodes. As a result, it can be concluded that the cost $\mathbf{W}_{0,0}$ at the root node (0,0) is equal to the sum of the costs of all leaf nodes.

It is worth noting that $\mathbf{W}_{0,0} = (\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})^{\mathrm{T}}$, is the square of the 2-norm of the prediction error, where $\hat{\mathbf{y}} = \mathbf{x}_{0,0}\{\breve{u}_{0,0}\} = \breve{\mathbf{u}}_{0,0}$ (see Fig. 7.3 for an example) is the model

estimation for the plant output $\mathbf{y}$. Since $\mathbf{W}_{0,0}$ is equal to the sum of the costs of all leaf nodes, it can be seen that, for a given tree structure, the 2-norm of the prediction error for the overall model can be minimized by adjusting each subband model separately.

Choice of the best wavelet packet tree

The choice of the best wavelet packet tree can be performed by applying the calculated cost mentioned in Subsec. 7.5.2. Based on the method of the choice of the best wavelet packet tree described in Section 6.4, for each node, we calculate the cost, the square of the 2-norm of the residue $\mathbf{e}_{i,j}$ and label this result $\mathbf{W}^1_{i,j}$. The superscript 1, which stands for a leaf node, takes the same meaning as the notation $*$ used in Section 6.4, and represents that $\mathbf{W}^1_{i,j}$ is equal to the cost $\mathbf{W}_{i,j}$ at node (i,j) only if this node is chosen to be a leaf. In order to decide if each node (i,j) should be a leaf or a non-leaf node, a comparison of the cost $\mathbf{W}^1_{i,j}$ (cost if (i,j) is a leaf node) with the sum $\mathbf{W}_{2i,j+1} + \mathbf{W}_{2i+1,j+1}$ of the costs at its child nodes (cost if (i,j) is a non-leaf node) is carried out using the decision rule,

$$\mathbf{W}_{i,j} = \begin{cases} \mathbf{W}^1_{i,j}, & \mathbf{W}^1_{i,j} \leq \rho(\mathbf{W}_{i,j+1} + \mathbf{W}_{2i+1,j+1}) \\ \mathbf{W}_{i,j+1} + \mathbf{W}_{2i+1,j+1} & \text{otherwise} \end{cases} \tag{7.62}$$

where the penalty factor ρ ensures that node (i,j) will only be split into child nodes if the cost reduction is large enough, aiming to justify the increase in model complexity. Fig. 7.5 illustrates this rule.

In the present case, the penalty factor is required to avoid an overfitting of the identification data (Coifman and Wickerhauser 1992). The choice of the penalty factor ρ is viewed as a model order determination problem, since the the number of nodes in the resulting tree is changed with the value of ρ. Generalized Cross Validation (GCV) techniques are adopted in the present work to address the determination problem and to solve the value of ρ. The techniques estimate the generalization performance of the model on the basis of the same signals used in its identification (Sjöberg *et al.* 1995), following the same approach adopted in (Zhang *et al.* 1997).

The GCV index ($\mathrm{GCV_{In}}$) for the wavelet packet model is defined as

$$\mathrm{GCV_{In}} = \mathrm{MSE} + \frac{2\rho}{N}\sigma_e^2 \tag{7.63}$$

where MSE is the mean-square error between the model output and the measured output, that is,

$$\mathrm{MSE} = \frac{1}{N}\Sigma_{k=1}^{N}(y_k - \hat{y}_k)^2 = \frac{1}{N}\mathbf{W}_{0,0} \tag{7.64}$$

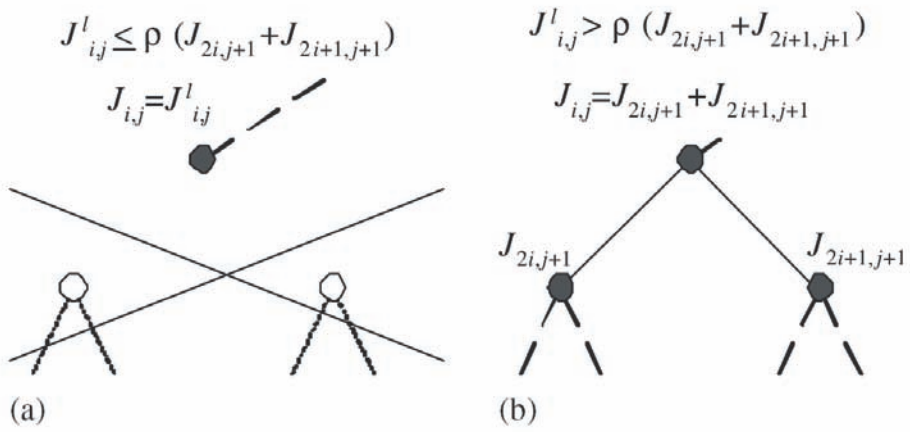


Figure 7.5. The Choice of the best decomposition tree. Decision if node (i, j) will be a (a) leaf or (b) non-leaf node. After Paiva and Galvão (2006).

where y_k and $\hat{y}_k$ denote elements of vectors $\mathbf{y}$ and $\hat{\mathbf{y}}$ (system and model outputs, respectively), N is the number of elements of such vectors and p is the number of adjustable parameters in the complete model. The measured output is assumed to be corrupted with an additive white noise of variance σ_e^2.

Since there is a subband model $M_{i,j}$ with three adjustable parameters ($s_{i,j}$, $\alpha_{i,j}$ and $\beta_{i,j}$) for each leaf node in the tree, it follows that p is equal to three times the number of leaf nodes. The number of leaf nodes in the tree depends on the penalty factor ρ. Therefore, the number of adjustable parameters p is a function of ρ.

The noise variance σ_e^2 is usually unknown and needs to be estimated. For the best value of ρ (the value that minimizes $\mathrm{GCV_{In}}$), the MSE between $\mathbf{y}$ and $\hat{\mathbf{y}}$ is a good approximation for σ_e^2. Therefore, from Eqs. (7.63) and (7.63), it follows that

$$\mathrm{GCV_{In}} = \mathrm{MSE} + \frac{2p}{N}\mathrm{MSE} = \frac{1}{N}\mathbf{W}_{0,0}\left(1 + \frac{2p}{N}\right) \tag{7.65}$$

in which both $\mathbf{W}_{0,0}$ and p depend on ρ.

Selecting the value of ρ that minimizes $\mathrm{GCV_{In}}$ provides a tradeoff between model parsimony and identification accuracy. Such a procedure improves the ability of the model to represent the behaviour of the system for input signals (Zhang *et al.* 1997).

7.5.3 System identification via optimized Mertz apodization functions

This subsection discusses use of optimized Mertz asymmetric triangular windows for system identification. The design of this window is a trade-off of between minimizing sidelobe levels in the calculated spectrum, which results in leakage of energy across different frequency bins and main lobe width, which leads to a reduction in frequency resolution.

The sample waveform is assumed to be the result of a discrete-time stochastic process $x[k]$, $k = 0, 1, ..., 2L-1$, composed of a true signal term $x_r[k]$ and a zero-mean, homoscedastic white-noise term $e[k]$ so that $x[k] = x_r[k] + e[k]$, where $x_r[k] = \mathbb{E}\{x[k]\}$, $\mathbb{E}\{e[k]\} = 0$, $\forall k$, $\mathbb{E}\{e[k_1]e[k_2]\} = 0$, $\forall t_1 \neq t_2$, and $\mathbb{E}\{e^2[k]\} = \sigma^2$, $\forall k$, with $\mathbb{E}$ denoting the expectation operator and σ the standard deviation of noise. The apodized waveform $x_a[k]$ is $x_a[k] = x[k]w[k]$, where $w[k]$ is an Mertz window to be optimized. We assume that $w[k]$ is a linear combination of m functions $\{g_i[k], i = 1, ..., m\}$, taken from a given library, that is

$$w[k] = \sum_{i=1}^{m} p_i g_i[k] \tag{7.66}$$

where $p_i, i = 1, ..., m$, are the decision variables for the optimization problem. The cost function adopted in this formulation is the expected value of the squared 2-norm of the error between $x_a[k]$ and the true signal term $x_r[k]$, that is,

$$\begin{aligned} W[\mathbf{p}] &= \sum_{k=0}^{2L-1} \mathbb{E}[x_a[k] - x_r[k]]^2 \\ &= \sum_{k=0}^{2L-1} x_r^2[k] + \sigma^2 w^2[k] - 2x_r^2[k]w[k] + x_r^2[k] \end{aligned} \tag{7.67}$$

where $\mathbf{P} = [p_1 p_2 \cdots p_m]^{\mathrm{T}}$. According to Eqs. (7.66) and (7.67), it follows that

$$W[\mathbf{p}] = 0.5\mathbf{p}^{\mathrm{T}}\mathbf{\Omega}\mathbf{p} + \mathbf{f}^{\mathrm{T}}\mathbf{p} + c \tag{7.68}$$

where

$$\begin{aligned} \mathbf{\Omega} &= 2\sum_{k=0}^{2L-1} x_r^2[k] + \sigma^2 \boldsymbol{Q}[k] \\ \mathbf{f} &= -2\sum_{k=0}^{2L-1} x_r^2[k]\boldsymbol{r}[k] \\ c &= \sum_{k=0}^{2L-1} x_r^2[k] \end{aligned} \tag{7.69}$$

with

$$\mathbf{Q}[k] = \begin{bmatrix} g_1^2[k] & g_2[k]g_1[k] & \cdots & g_m[k]g_1[k] \\ g_1[k]g_2[k] & g_2^2[k] & \cdots & g_m[k]g_2[k] \\ \vdots & \vdots & \cdots & \vdots \\ g_1[k]g_m[k] & g_2[k]g_m[k] & \cdots & g_m^2[k] \end{bmatrix} \tag{7.70}$$

and

$$\mathbf{r}[k] = [g_1[k] \; g_2[k] \; \cdots \; g_m[k]]. \tag{7.71}$$

Since the true signal term $x_r[k]$ is unknown, the actual measurement $x[k]$ for a particular realisation of the stochastic process can be used as an estimate $\hat{x}_r[k]$ for $x_r[k]$. An estimate $\hat{\sigma}$ for σ can be calculated as the standard deviation of the early part of the measured time-domain signal before the arrival time of the pulse.

Without any constraints to be placed in the optimization process, the minimum of $W[\mathbf{p}]$ would be obtained from the equation $\nabla W[\mathbf{p}] = 0$, which leads to $\mathbf{p} = -\mathbf{\Omega}^{-1}\mathbf{f}$ provided that $\mathbf{\Omega}$ is nonsingular. However, constraints such as nonnegativity of the elements of the windowing function $(w[k] \geq 0)$, monotonicity $(\mathrm{d}w[k]/\mathrm{d}k \leq 0)$, and endpoint restrictions $(w[0] = 1,\ w[2L-1] = 0)$ must be imposed on the optimization process to ensure that the function is well behaved in both time and frequency domains. To satisfy the mentioned constraint conditions, the Mertz asymmetric triangular window is designed with a maximum at the point corresponding to the absolute maximum of the waveform and falls to zero at the borders.

The endpoint constraints $w(0) = 1$ and $w(2L-1) = 0$ can be translated as the following linear equality restrictions on the coefficient vector $\mathbf{p}$:

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & g_2[2L-1] & g_3[2L-1] & \cdots & g_m[2L-1] \end{bmatrix} \mathbf{p} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{7.72}$$

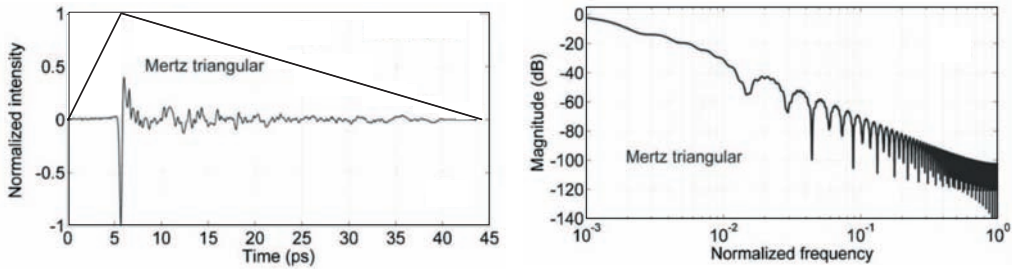
The positivity constraint for the Mertz window can be translated as

$$p_i \geq 0, i = 2, \cdots, m. \tag{7.73}$$

The first coefficient p_1, which is associated with a constant term, is free to assume negative values to satisfy the endpoint constraints described in Eq. (7.72). The minimization of the cost function $W[\mathbf{p}]$ given in Eq. (7.68) subjects to the constraints in Eqs. (7.72) and

(7.73) is a quadratic programming problem, which can be solved by using efficient numerical algorithms.

Fig. 7.6 shows a THz waveform with application of conventional the Mertz asymmetric triangular window function.



(a) (b)

Figure 7.6. Plots of Mertz asymmetric triangular windows on the measured THz signals. (a) THz waveform with application of conventional Mertz asymmetric triangular windows. (b) Corresponding frequency domain plots of Mertz windowing functions (low-frequency value normalized to 0 dB). After Galvão *et al.* (2007).

7.6 Chapter summary

This Chapter investigates different feature extraction methods to achieve the transformed signal characteristics of THz measurements. It is important for pattern recognition tasks to convert patterns to features and to describe the collected data in a compact form. A series of signal processing techniques have been adopted to extract feature sets. The aim is to explore the ability of these extracted features to reveal important differences for discrimination between patterns associated with different classes. These feature extraction methods are summarised as follows:

(i) Fourier transform for signal analysis has been the standard analytical tool to be adopted by engineers and scientists, which defines a relationship between a signal in the time domain and its representation in the frequency domain.

(ii) Compared to Fourier basis functions with infinite support in time, wavelet transforms afford good time-frequency localization features. Also AR and ARMA models

on the wavelet transforms of T-ray pulse responses have been used as feature extraction methods for classification purposes. The correlation and Durbin's methods are optimal for the calculation of the AR and ARMA model parameters.

(iii) A subspace system identification algorithm as one of feature extraction algorithms is based on geometric operations on subspaces spanned by matrices obtained from measurements. Despite not using any explicit optimization criteria, such as in ARMA processing, subspace methods are fast and reliable.

(iv) A subband model identification algorithm via the choice of best wavelet packet tree holes great potential, especially for more spectrally rich samples, since it supplies spectrum information with respect to the both space localisation and frequency localisaton. Additionally, the best wavelet packet tree choice provides a tradeoff between model parsimony and identification accuracy.

(v) System identification via optimized Mertz apodization functions as a feature extraction method is also discussed. The tailored design of the Mertz window is specifically applied to the asymmetric THz transient waveforms. It aims to afford more algorithms for comparison with other signal processing techniques. Since pattern recognition tasks are problem dependent, these feature extraction algorithms play an important role in exploration of THz signal recognition for different biomedical specimens.

The last phase of the current THz pattern identification system is introduced in the next Chapter, which describes different pattern classification schemes, falling within a supervised learning strategy. These classification schemes are generally validated via several different evaluation approaches, with aims to achieve the improvements on classification components.

Chapter 8

Pattern Classification

IN signal processing, pattern classification refers to the separation of patterns, measured or observed, into small classes, and then the assignment of each new existing pattern to a particular class based on the knowledge of the patterns. Classifying or describing observations is realised via a classification or description scheme. The classification or description scheme is usually based on training sets that have already been classified or described. This is called a supervised learning strategy. Learning can also be unsupervised, but this is not within the scope of this Thesis. The classification or description scheme in this Thesis is mainly concerned with statistical (or decision-theoretic) approaches. They include the Mahalanobis distance classifier, the Euclidean discrimination matrix, and Support Vector Machines (SVMs).

8.1 Mahalanobis distance

In this Section, the Mahalanobis distance classifier is introduced. An application of this classifier to the extracted features using ARMA models over wavelet subbands is provided. The other experimental instances regarding applying Mahalanobis distance classifiers on various THz measurements to evaluate classification performance of different feature extraction techniques will be represented in the next Chapter.

8.1.1 Definition of Mahalanobis distance

One of the most common and versatile classifiers is the Mahalanobis distance classifier (Schürmann 1996). It is a type of minimum distance classifier. It assumes that the data for each class are normally distributed, thus the samples, $\mathbf{x}_m$, drawn from each class will form a cluster in k dimensions, with a centre given by the mean vector, $\mathcal{A}_m$, calculated by the expected value $\mathbb{E}$ of the m^{th} entry in the vector $\mathbf{x}$, and shape dependent on the covariance matrix, $\mathbf{C}$. Estimates are formed for these parameters regarding each class m, using the training vectors,

$$\mathcal{A}_m = \mathbb{E}[\mathbf{x}_m] \tag{8.1}$$

$$\mathbf{C}_m = \mathbb{E}[(\mathbf{x}_m - \mathcal{A}_m)(\mathbf{x}_m - \mathcal{A}_m)^{\mathrm{T}}]. \tag{8.2}$$

The *Mahalanobis distance* calculates the distance of a given vector $\mathbf{x}$ from the mean vector $\mathcal{A}_m$ for a given class normalised by the variance $\mathbf{C}_m$ of training vectors in that direction. For a given class, m, the distance is calculated by,

$$\rho_m(x) = \sqrt{(\mathbf{x} - \mathcal{A}_m)^{\mathrm{T}} \mathbf{C}_m^{-1} (\mathbf{x} - \mathcal{A}_m)}. \tag{8.3}$$

Classification is then performed by assigning a label to the given vector for which the Mahalanobis distance is minimized. This classifier is illustrated as Fig. 8.1

The Mahalanobis-based classification scheme is optimal for normally distributed classes with equal covariance matrices (linear discriminant) and equal *a priori* probabilities. This Mahalanobis-based classifier is chosen because it is simple to implement and it provides reasonable results for a variety of statistical properties, thereby highlighting the performance of the associated feature extraction techniques. More complicated classification algorithms abound and the appropriate choice for THz applications is an open research area.

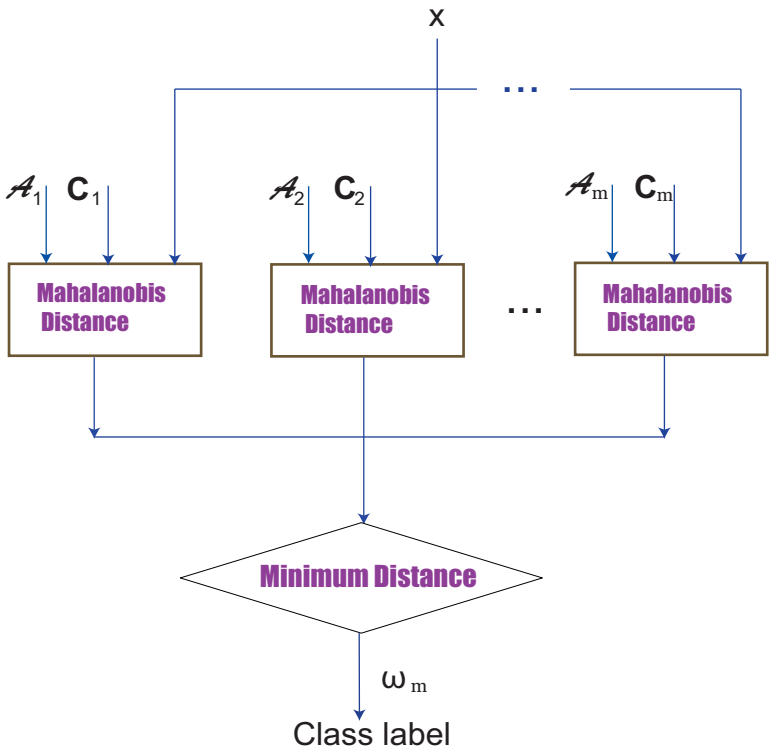


Figure 8.1. A schematic of the Mahalanobis distance classifier. The *Mahalanobis distance* of a given vector $\mathbf{x}$ from the mean vector $\mathcal{A}_m$ for a given class is normalised by the variance $\mathbf{C}_m$ of training vectors in that direction.

8.1.2 The Mahalanobis classifier via extracted features

For a practical application consideration, we formulate the Mahalanobis classification scheme on a set of feature matrices of ARMA modeling over several variable wavelet subbands, as described in Section 7.4. The corresponding classifier design based on the feature extracted scheme is expressed herein. For a given class, m, the distance from a feature matrix $\mathbf{DC}_{\mathbf{j}}^{l}$ to the class mean $\mathcal{A}_m$, is defined as

$$\rho_m(X) = \sqrt{(\mathbf{DC}_{\mathbf{j}}^{l} - \mathcal{A}_m)^{\mathrm{T}} \mathbf{C}_m^{-1} (\mathbf{DC}_{\mathbf{j}}^{l} - \mathcal{A}_m)} \tag{8.4}$$

where $\mathbf{C}_m$ is the covariance matrix of the feature vectors regarding class m, $\mathbf{DC}_{\mathbf{j}}^{l}$ with $l = 1, 2, 3$ represents the averaged coefficients matrix related to AR ($l = 1$), MA ($l = 2$), and ARMA ($l = 3$) modeling of wavelet approximation coefficients at three decomposition levels $\mathbf{j}$, that is, $\mathbf{DC}_{\mathbf{j}}^{1}$ being $\mathbf{DC}_{\mathbf{j}}^{\mathrm{AR}}$, $\mathbf{DC}_{\mathbf{j}}^{2}$ being $\mathbf{DC}_{\mathbf{j}}^{\mathrm{MA}}$, $\mathbf{DC}_{\mathbf{j}}^{3}$ being $\mathbf{DC}_{\mathbf{j}}^{\mathrm{ARMA}}$. In practice, the covariance matrix is estimated from the training vectors. During classification, the minimum Mahalanobis distance from feature matrix $\mathbf{DC}_{\mathbf{j}}^{l}$ to each class centre $\mathcal{A}_m$ is used to assign the appropriate class label.

8.2 The Euclidean discrimination metric

In what follows, the Euclidean discrimination metric is presented, which is applied to evaluate how well multi classes can be distinguished within the THz frequency range under consideration, illustrated in Section 9.3. This method is similar to the Mahalonobis discrimination. Its difference from Mahalanobis distance lies in that the Euclidean discrimination is the measurement of all variables in the same units of length (Mullen and Ennis 1987, McLachlan 2005). The Euclidean discrimination metric is described as follows, with link to the current T-ray project.

Let $x_{m,n}$ be the THz frequency response of the m^{th} object at the n^{th} spectral bin. A row vector x_m is defined for each object in the form:

$$\mathbf{x}_m = [x_{m,1} \;\; x_{m,2} \cdots x_{m,n}]. \tag{8.5}$$

Let $\mathcal{A}_1, \mathcal{A}_2, ..., \mathcal{A}_\varsigma$ denote the mean value of the objects belonging to classes $1, 2, ..., \varsigma$ respectively, that is:

$$\mathcal{A}_\varsigma = \frac{1}{N_\varsigma} \sum_{m \in I_\varsigma} \mathbf{x}_m \tag{8.6}$$

where I_ς is the index sets of objects belonging to class ς, and N_ς is the number of objects in each class. A between-class dispersion metric is calculated as:

$$\mathbf{D}_B = \frac{1}{\varsigma} \sum_{\omega=1}^{\varsigma} ||\mathcal{A}_\omega - \overline{\mathcal{A}}||^2 \tag{8.7}$$

where

$$\overline{\mathcal{A}} = \frac{1}{\varsigma}(\mathcal{A}_1 + \mathcal{A}_2 + \cdots + \mathcal{A}_\varsigma). \tag{8.8}$$

A within-class dispersion metric is calculated for each class ω as:

$$\mathbf{D}_{W,\omega} = \frac{1}{N_\omega} \sum_{m \in I_\omega} ||\mathbf{x}_m - \mathcal{A}_\omega||. \tag{8.9}$$

An overall within-class dispersion metric $\mathbf{D}_W$ is calculated as:

$$\mathbf{D}_W = \frac{1}{\varsigma}(\mathbf{D}_{W,1} + \mathbf{D}_{W,2} + ... + \mathbf{D}_{W,\varsigma}). \tag{8.10}$$

Finally, the discrimination metric $\mathbf{F}$ is defined as:

$$\mathbf{F} = \frac{\mathbf{D}_B}{\mathbf{D}_W}. \tag{8.11}$$

It is not difficult to see that the smaller value of within-class dispersion metric $\mathbf{D}_W$, the larger value of the discrimination metric $\mathbf{F}$. It expects to achieve a larger value of resultant discrimination matric $\mathbf{F}$, so as to obtain well discrimination of data points from various classes.

8.3 Overview of SVMs

Recent advances in statistical learning theory (Vapnik 1995, Vapnik 1998) have made the Support Vector Machine (SVM) a popular machine learning algorithm. This algorithm maps data into a high-dimensional feature space, then a separating hyperplane is established, which maximizes the boundary margin between two classes. At its core, a SVM is a two-class classifier. With the recent improvements, SVMs have been extended to solve multiclass classification problems. The following subsections describe each in turn.

8.3.1 Binary classification of SVMs

Support vector machines are binary classifiers, which classify data depending on a set of support vectors (Cristianini and Shawe-Taylor 2000). The training data sets come from a set of labelled samples called learning vectors. We denote such a set of learning vectors as $(\mathbf{x_i}, y_i) \in R^N \times \{\pm 1\}, i = 1, ..., l$, and y_i denote the class label corresponding to each input vector $\mathbf{x}_i$. The support vectors are subsets of the training data sets and are used to construct a l dimensional hyperplane in feature space, which acts as a boundary separating the different classes. A decision function $f(\alpha) : R^N \rightarrow \pm 1$ is achieved based on a given class function $f(\alpha) : \alpha \in \Lambda$, with the aim to correctly assign class labels to test samples $\mathbf{x}$. The Vapnik-Chervonenkis (VC) dimension (Vapnik 1998, Vapnik 1995, Muller *et al.* 2001) is a property of a set of functions $f(\alpha)$, which is defined as the maximum number of training points that can be segmented by $f(\alpha)$. Note that α corresponds to the weights and biases, which can be adjusted to label the output $f(\mathbf{x}, \alpha)$ based on the input $\mathbf{x}$. The expectation of the test error for a learning machine is:

$$R(\alpha) = \int \frac{1}{2}|y - f(\mathbf{x}, \alpha)| \mathrm{d}P(\mathbf{x}, y) \tag{8.12}$$

where, $R(\alpha)$ is called the expected risk. It is the quantity connected with density $p(\mathbf{x}, y)$ that we are ultimately interested in. The 'empirical risk' $R_{\mathrm{emp}}(\alpha)$ is defined to be the measured mean training error for a fixed, finite number of observations:

$$R_{\mathrm{emp}}(\alpha) = \frac{1}{2l} \sum_{i=1}^{l} |y_i - f(\mathbf{x}_i, \alpha)|. \tag{8.13}$$

The quantity $\frac{1}{2}|y_i - f(\mathbf{x}_i, \alpha)|$, which is called the loss, takes on values equal to 0 or 1. When we set the probability to $1 - \eta$, we achieve the following bound:

$$R(\alpha) \leq R_{\text{emp}}(\alpha) + \sqrt{\left(\frac{h(\log(2l/h)+1) - \log(\eta/4)}{l}\right)}. \tag{8.14}$$

The non-negative integer h is called the VC dimension, and allows the measure of the *capacity* that is suitable for the amount of available training data. When a sufficiently small η is selected, the right hand side of the equation is minimized, the $f(\mathbf{x}, \alpha)$ functions give the lowest upper bound of the actual risk. It is the basic idea of structural risk minimization.

Since minimizing the training error (the computation of VC-dimension) does not guarantee a small test error, in order to make the decision function f perform well on unseen pattern, the principle of structural risk minimization needs to be applied to minimize test error and achieve a capacity that is suitable for the amount of available training data sets.

The learning algorithm is designed to allow the support vectors computed via performing structure risk minimization. In other words, a VC-dimension bound is calculated to identify the optimal hyperplane and maximizes the margin of the nearest learning vectors. The decision hyperplane is calculated based on the following equation:

$$f(\mathbf{x}) = \text{sgn}(\sum_{i=1}^{s} y_i \alpha_i (\mathbf{x}_i \cdot \mathbf{x}) + b) \tag{8.15}$$

where $\mathbf{x}_i, (i = 1, ..., s)$ are support vectors, which are the closest points from the training vectors (learning vectors) to the separate hyperplane and sgn is called the signum function. The solution of this large-scale quadratic programming problem is applied to calculate the coefficients α_i and b. The procedure is to solve the dual problem, which is to maximize

$$L(\alpha) = \sum_{i=1}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j (\mathbf{x}_i \cdot \mathbf{x}_j) \tag{8.16}$$

subject to $\sum_{i=1}^{l} \alpha_i y_i = 0$ and $0 \leq \alpha_i \leq C$ for $i = 1, ..., l$.

The penalty parameter C is selected by the user, which is viewed as a regularization parameter for the linearly inseparable learning vectors aiming to accept the possible misclassifications.

Note that SVMs use a kernel function (Schölkopf and Smola 2002, Shawe-Taylor and Cristianini 2004), which allows fitting the hyperplane to the data. Instead of a linear classifier, which is limited to producing linear decision surfaces, the hyperplane

(Hearst 1998) needs to be augmented to fit the nonlinear decision surfaces. A dot product space is constructed by mapping data, which is realised by performing a nonlinear map $\phi : R^N \rightarrow F$. The above linear algorithm then can be applied in the feature space F. The solution satisfies the following function

$$f(\mathbf{x}) = \text{sgn}(\sum_{i=1}^{s} y_i \alpha_i \phi(\mathbf{x}_i) \cdot \phi(\mathbf{x}) + b). \tag{8.17}$$

This is a nonlinear function of the original input vectors $\mathbf{x}$.

According to Cover's theorem (Cover 1965), a new feature space F can be achieved via transforming a multidimensional space, where the dimensionality of the feature space is high enough to allow the target patterns to be linearly separable with a high probability. The inner products (dot products) enable the high dimensional space to be treated easily, which are indicated by the ϕ mapping. Accordingly, the kernel function K is defined as:

$$K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}) \cdot \phi(\mathbf{y}). \tag{8.18}$$

The kernel production substitutes for all the occurrences of a dot product resulting from two mappings.

There are four popular kernel functions:

- linear Kernel:

$$K(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\text{T}}\mathbf{y} \tag{8.19}$$

- polynomial kernel:

$$K(\mathbf{x}, \mathbf{y}) = (\gamma \mathbf{x}^{\text{T}}\mathbf{y} + r)^p, \gamma > 0 \tag{8.20}$$

- RBF (Gaussian) kernel:

$$K(\mathbf{x}, \mathbf{y}) = \exp(-\lambda \left\|\mathbf{x} - \mathbf{y}\right\|^2), \lambda > 0 \tag{8.21}$$

- Hyperbolic tangent kernel:

$$K(\mathbf{x}, \mathbf{y}) = \tanh(\gamma \mathbf{x}^{\text{T}}\mathbf{y} + r), \gamma > 0 \tag{8.22}$$

where $\mathbf{x}, \mathbf{y}$ are SVM data vectors, T labels vector transpose, γ and r label the scale and offset of the corresponding kernels (they are normally set to 1), p labels the degree of polynomial kernel, and λ labels the width parameter of Gaussian kernel.

In the instance of identifying RNA samples, presented in Section 9.5, the RBF kernel function is proposed as a choice, and it is found to give good classification performance. Meanwhile, it is found that polynomial kernels are best suited to achieve multiclass classification of powdered samples. Linear and hyperbolic tangent kernels are not employed since they have been found to be unsuitable for the T-ray data set.

8.3.2 Pairwise SVM classification

The previous Section described a SVM for two-class pulsed signal classification, which is also called a *dichotomy*. This is appropriate for the object detection application of two-class T-ray pulses and discriminating an object from the background. However, the majority of object recognition problems consist of more than two substances. Consequently, extended SVMs are required for application to multiclass pulsed signals. The optimal design of multiclass SVM classifier is an area of active research. One frequently used method is a pairwise classifier, based on one-against-one decomposition (Schölkopf *et al.* 1998), which works by using a decision function f_{kl}. Here, kl indicates each pair of classes selected from separated target classes. Since the symmetric property of pairwise approach, it holds that $f_{kl} = -f_{kl}$, where f_{kl} satisfies the following equation:

$$f_{kl}(\mathbf{x}) = \mathbf{w}_{kl}\mathbf{x} + b_{kl} \tag{8.23}$$

where $\mathbf{w}$ is normal to the hyperplane between class k and class l, $|b|/||\mathbf{w}||$ is the perpendicular distance from the hyperplane to the origin, and $||\mathbf{w}||$ is the Euclidean normal of $\mathbf{w}$ with a vector dimension of M.

The signum function is used for the hard threshold decisions:

$$\operatorname{sgn}(f_{kl}(\mathbf{x})) = \begin{cases} 1, & f_{kl}(\mathbf{x}) > 0 \\ -1, & f_{kl}(\mathbf{x}) \leq 0. \end{cases} \tag{8.24}$$

The class decision can be achieved by summing up the according pairwise decision functions:

$$f_k(\mathbf{x}) = \sum_{k \neq l, l=1}^{n} \operatorname{sgn}(f_{kl}(\mathbf{x})) \tag{8.25}$$

where n is the number of the separated target classes.

The pairwise classifier proceeds as follows: assign a label to the class, $\{\arg\ \max\ f_k(\mathbf{x}), (k = 1, ..., n)\}$. The max number of votes for k classes holds $\{\max\ k \rightarrow f_k = (k -$

1)}. If Eq. (8.25) is satisfied for $\{\max k \rightarrow f_k < (k-1)\}$, the $\mathbf{x}$ is unclassifiable. The pairwise classification converts the n-class classification problem into $n(n-1)/2$ two-class problems which cover all pairs of classes.

The procedure for pairwise classification is illustrated in Fig. 8.2.

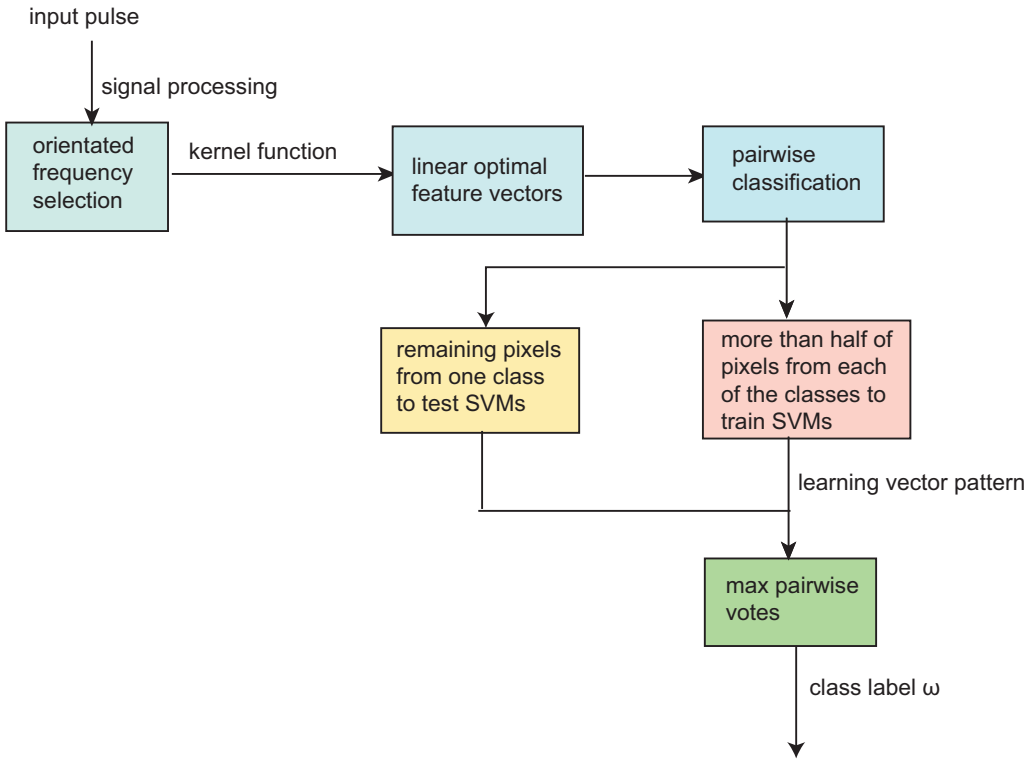


Figure 8.2. Illustration of the procedure for pairwise classification.

8.4 Classifier design

An important aspect of pattern recognition is the design of the classifier. This is because perfect classification performance is often impossible, and therefore a more general task is to determine the probability for each of possible classes. The probability is provided by feature-vector representation of the input data, which allows a largely domain-independent theory of classification (Duda *et al.* 2001).

Due to the complexity of a classification system and sources of random noise in THz measure systems, some variability exists among feature values for objects in the same category. This results in difficulties in finding the correlation between the variability in the feature values from the same category and the difference between feature values

for objects extracted from different categories. Therefore, it is important to find the best way to design a classifier via extracted features, in order to cope with this variability and to achieve the best classification performance. This is known more simply as 'robustness'.

8.4.1 Learning and adaption

To design or create a classifier, there are two steps: positing some general form of model, i.e. the network size and learning parameters, with input to some form of classifier, and then using training patterns to learn or estimate the classifier. In general, training the classifier refers to the process of using data to determine the classifier's parameters. Learning is employed in the stage of training samples for design of a classifier. It refers to some form of algorithms for reducing the error on a set of training data. Since it is difficult to guess the best classification decision ahead of time, it is common to expend great time to achieve proper learning. In this Thesis, supervised learning, as one of many different procedures for training classifiers and choosing models, is used exclusively. In a supervised algorithm, a class label for each pattern in a training set is provided, with aims to transform the pattern data in testing sets into a series of different target classes. The Thesis for pattern recognition proves that the supervised learning algorithm is powerful to learn the solution to a given problem and it is stable to parameter variations in a THz frequency range. Fig. 8.3 illustrates the design of a classifier for a pattern recognition procedure. A set of imaging data x_i is automatically transformed into a series of classes ω_ς.

8.4.2 Evaluation

Evaluation implies measurement of the performance of a classification system and identification of the need for improvements on classification components (Duda *et al.* 2001). There are different ways to evaluate a classifier, including different types of cross-validation methods, the use of measure functions, and graphical analysis of lift charts and Receiver Operating Characteristic (ROC) curves, etc. The choice of which one to use depends on many attributes and, according to Mitchell (1997), there is no method that satisfies all the constraint conditions. This subsection mainly talks about former two evaluation methods, both of which are frequently used in this Thesis.

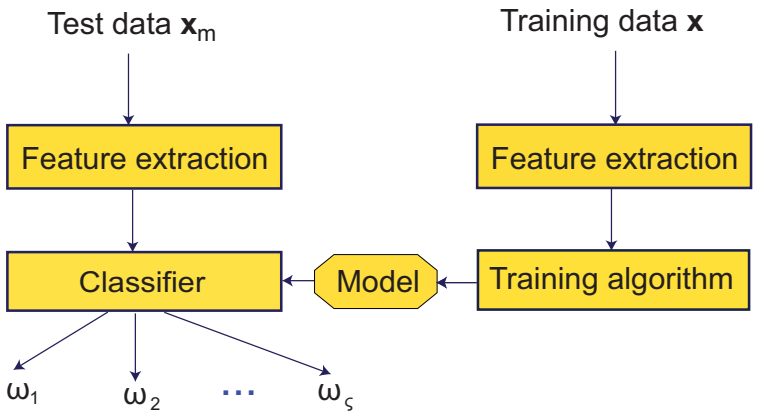


Figure 8.3. Design of a classifier. A set of imaging data x_m is automatically transformed into a series of classes ω_ς. This procedure is realised via positing some general form of model, and then training patterns is input to learn or estimate the unknown parameter model. The test data are subsequently evaluated based on the known training sets, the result of which validates the classifier.

Cross-validation is a statistical technique widely used for estimating generalization error based on 'resampling' (Weiss and Kulikowski 1991, Shao 1993, Kearns 1997). The procedure is to divide the available data into a number of folds or partitions. Training is then performed on all except one partition, which is left for testing. There are three common types for cross-validation.

The holdout method is the simplest kind of cross-validation. The data set is separated into two sets, called the training sets and the testing sets. In this method, the training data are used to predict or estimate the output values of the test data. The errors it makes are accumulated to give the mean absolute test set error, which is used to evaluate the model. This method is usually preferable to a residual method and takes no longer to compute. However, its evaluation can have a high variance. The evaluation may depend heavily on which data points are used in the training sets, and which are used in the test sets, and thus the evaluation may be significantly different depending on how the division is made. In this Thesis, we propose to iteratively apply the holdout method via halving the whole data sets with different compositions until a large number of different sets are achieved, and then calculate the average value across all runs of the holdout tests. The aim is to reduce the variation of resultant evaluation. This evaluation method is used in the next Chapter in several THz discrimination experiments.

K-fold cross-validation is one way to improve evaluation accuracy over the holdout method. The data set of N points can be partitioned into K sets ('folds') with size N/K for a K-fold cross-validation. In this issue, the holdout method is repeated K times. Each time, one of the K subsets is used as the test set and the other $K-1$ subsets are put together to form a training set. Then the average error across all K trials is computed. The advantage of this method is that it is affected less by the division of the data. Every data point appears in a test set exactly once, and appears in a training set $K-1$ times. The variance of the resulting estimate is reduced with increased K. The disadvantage of this method is that it takes K times as much computation to make an evaluation, as the training algorithm has to be re-run K times for an ultimate evaluation.

Leave-one-out (LOO) cross-validation is K-fold cross-validation, with K equal to N, the number of data points in the data set. That means that for N times, the classifier is trained on all the data except for one point and a prediction is made for that point. This procedure is then performed until all points have been left out once. As before, the average error is computed and used to evaluate the model. The evaluation given by leave-one-out cross-validation error is good, but it is expensive to compute. More methods like locally weighted learners simplify LOO validation, and further details can be found in Atkeson *et al.* (1997).

In addition to the cross-validation, using measure functions (or metrics) as an alternative way of selecting and evaluating learned classifiers is proposed by Atkeson *et al.* (1997), which allows one to define the learning problem as a computational problem. The measure function, such as Euclidean discrimination, takes three aspects of the classifier into account: (i) subset fitting, (ii) similarity, and (iii) simplicity. Subset fitting is the most commonly used measurement of classifier performance. Subset fitting measures classification correctness regarding known data points. It is evaluated via performing the cross-validation (CV) method, and then an average value on the resultant evaluations is calculated, which is helpful in tuning or choosing algorithms. Similarity measures the discrimination capabilities of a classifier. Similar test data points should then be clustered similarly. It should be expressed in terms of distances between training instances and decision borders. That is, the correctly classified points should preferably reside at 'safe' distances from decision borders. Because of the increased complexity with the number of data attributes, the partitioning of the universe should be as simple as possible, with the proposition of the good classification accuracy. This

is related to simplicity. The measure metric (function) is applied in Section 9.3 to evaluate the ability of system identification to extract the significant features from THz measurements.

8.4.3 Overfitting

The problem of 'overfitting' appears where a complex classifier, such as neural networks, that allows perfect classification of the training sample, fails to perform well on new patterns (Duda *et al.* 2001). It is an important area of research in statistical pattern classification to determine how to adjust the model complexity. One possibility is to reduce the dimensionality, either by redesigning the feature extractor, by selecting an appropriate subset of the existing features, or by combing the existing features. Another possibility is to assume that all ς classes share the same covariance matrix, and to cluster the available data together. More details are in the book by Duda *et al.* (2001).

The current classification schemes represented in this Thesis are relatively robust to overfitting problems. Since Mahalanobis distance displays near-optimum properties for a wide class of input data, it does not require fine tuning of classifier parameters, which reduces the risk of overfitting in classification problems (Ferguson 2004). In SVM learning, a kernel function plays the role of the dot product in the feature space. It allows a support vector machine to locate a separating hyperplane in the feature space and classify points in that space without ever representing the space explicitly. This technique avoids the computational burden of explicitly representing the feature vectors (Vapnik 1998).

8.5 Chapter summary

This Chapter describes different pattern classification schemes, which fall within a supervised learning strategy. These classification schemes are generally validated via several different evaluation approaches that aim to improve on classification steps.

The Mahalanobis distance classifier is a quadratic classifier, and as the most common and versatile classifiers it is simple to implement. The Mahalanobis distance takes into account the covariance among the variables in calculating distances. It is a very useful way of determining similarity of an unknown sample set to a known one. In contrast to the Mahalanobis distance, the Euclidean discrimination metric measures all variables

in the same units of length, and linearly evaluates how well the multiple classes can be distinguished. Support vector machines as a recent advance have become a popular machine learning algorithm in statistical learning theory. SVMs map input vectors to a higher dimensional space where a maximal separating hyperplane, or a maximal boundary margin between two classes is constructed. With the recent improvements, SVMs have been extended to solve multiclass classification problems. Therefore, in addition to a dual classification, this Chapter also discusses a pairwise extension strategy for SVMs.

Classifier design is an important strategy to achieve pattern recognition. A supervised algorithm is adopted, which needs to provide a class label for each pattern to a given test set via applying training data. There are several evaluation algorithms to measure the performance of these classification systems. Cross-validation is a statistical technique widely used for estimating generalization error, which consists of the holdout method, K-fold cross validation, and leave-one-out cross validation. In addition, measure functions (or matrices) as an alternative way of selecting and evaluating learned classifiers are also adopted in the Thesis for THz signal recognition. All of these classification schemes and evaluation algorithms will be experimentally illustrated in the next Chapter by applying them to a series of THz measurements.

Overfitting is a problem induced during pattern classification is performed. Since there is no requirement to finely tune Mahalanobis classifier parameters and to explicitly represent the feature space that SVMs involve, neither Mahalanobis distance nor SVMs are likely to be at risk of falling into the overfitting problem.

Feature extraction and classification algorithms described in these recent two Chapters are validated by using specific THz experiments. The next Chapter represents several important THz experiments and contains several central results of the relevant THz pattern recognition.

Chapter 9

THz Pattern Recognition Experiments

THE work described in this Chapter combines the different techniques discussed in previous Chapters into case studies for the biomedical specimen identification. Section 9.1 outlines the THz experiment setup. In Section 9.2, an improvement in classification accuracy is demonstrated by applying wavelet-based techniques in the preprocessing of T-ray pulsed signals. In Section 9.3, three system identification schemes for discriminating between lactose, mandelic acid, and dl-mandelic acid THz transients is proposed, with application of a discrimination metric for the evaluation of classification performance. Section 9.4 represents the implication of Auto Regressive (AR) and Auto Regressive Moving Average (ARMA) models on the wavelet transforms of measured T-ray pulse data for automatic classification of THz measurements, highlighting their potential in biomedical, pharmaceutical and security applications. Section 9.5 illustrates that support vector machine (SVM) learning algorithms are sufficiently powerful to detect patterns within noisy biomedical measurements. Case studies show effective discrimination of RNA samples and various powdered substances.

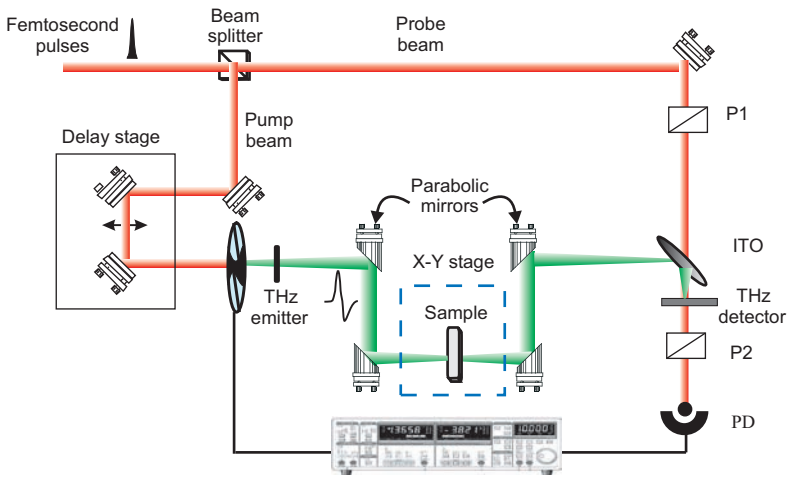


Figure 9.1. Illustration of a femtosecond laser-based T-ray functional imaging system. This T-ray functional imaging system is based on a pump-probe configuration. PD=infrared photodiode; P1,P2=polarizers; ITO=indium tin oxide coated beam splitter.

9.1 THz spectroscopy and imaging

T-ray time domain techniques allow the measurement of T-ray pulsed responses in both amplitude and phase (time delay). The T-ray time domain response is obtained via terahertz pulsed imaging (TPI).

Fig. 9.1(a) illustrates a schematic for a THz-TDS experiment setup for transmission-type terahertz pulsed imaging (TPI) (Ferguson *et al.* 2003). A pair of electrooptic generator and detector crystals are employed. A typical TPI system consists of an ultra-fast laser, an optical delay stage, a chopper, two optical rectification crystals, a photo detector, and paraboloidal mirrors. An object sample can be moved for imaging in the plane perpendicular to T-ray beam by an *x*-*y* computer-controlled translation stage. A synchronous sampling technique is used to detect the transmission terahertz radiation through the sample with collinear propagation between the pump and probe beams. The T-ray imaging system recovers full amplitude and phase information due to the ability to progressively alter the time delay between the pump and probe beams via the optical delay stage indicated in Fig. 9.1.

9.2 Enhanced T-ray signal classification using wavelet preprocessing

This Section performs a case study of one dimensional discrete wavelet transforms in the classification of T-ray pulsed signals. In our scheme, Fast Fourier Transforms (FFTs) are used as a feature extraction tool and a Mahalanobis distance classifier is employed for classification. Soft threshold wavelet shrinkage de-noising is then used and plays an important role in de-noising and reconstruction of T-ray pulsed signals. An iterative algorithm is applied to obtain three optimal frequency components and to achieve good classification performance.

In this Section, we discuss the ability of the wavelet techniques to assist in effective classification tasks, using T-ray measurement on a series of powder samples to illustrate our approach. It contributes to the field by demonstrating enhanced classification of T-ray pulsed signals via the application of wavelet-based preprocessing techniques.

9.2.1 Feature extraction

Feature extraction is an important step in all classification problems. The objective of feature extraction is to obtain the critical features from the T-ray signals to facilitate good classification performance. In this Thesis, the input features to the classifier are the heuristic SURE soft threshold shrinkage denoised amplitude and deconvolved phase frequency coefficients from the seven classes of T-ray data for different samples. The deconvolution is realised via dividing the sample spectral response from T-ray reference pulses by the system frequency response, in order to isolate system and ambient noise. To reduce the dimensionality and to simplify the computations, an iterative algorithm is proposed to identify a subset of the three available frequency components, which offer optimal classification accuracy. We used frequencies up to 1.5 THz in order to keep linear phases for the extracted features. Fig 9.2(a) shows several measured waveforms and Fig. 9.2(b) shows the phase plot corresponding to one of those waveforms.

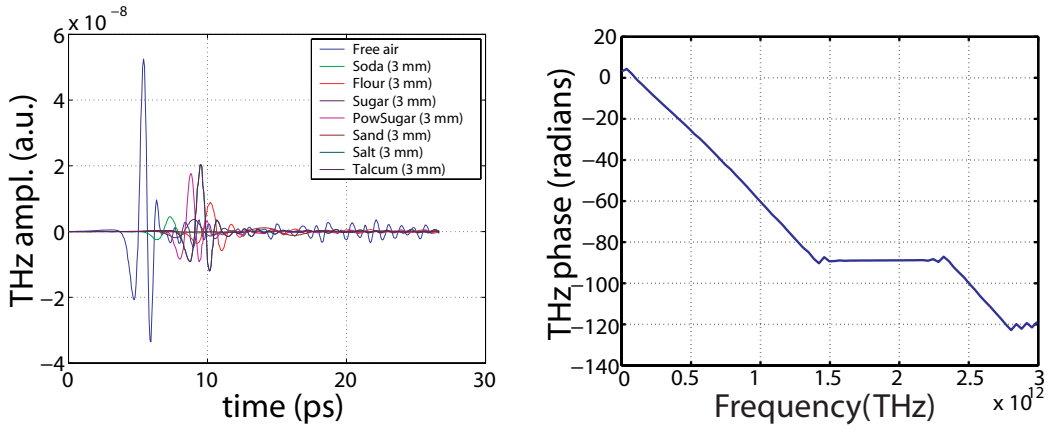
9.2.2 Classification

An important consideration in real-world detection problems is to test the ability of the classifier to accurately classify powders at the different sample thicknesses at which the classifier is trained.

In this Section, the classifier used is chosen to be a Mahalanobis distance classifier (Schürmann 1996). To validate the effectiveness of wavelet shrinkage de-noising, leading to the enhanced ability of the classifier to classify powders, several different powder samples are tested. The current powder experiment is to detect powders inside envelopes, where six different powder samples are taped onto a sheet of Thesis and then put inside an envelope. The traditional scanning imaging system is used to obtain a 2D THz image of the sample. This imaging system allows differing scattering paths and minor variations in powder thickness and density to be observed. Typically, a 1D image is sufficient (i.e. raster scan in only one direction), and a 51 pixel image (with a 100 μm spatial interval for one pixel) is acquired. For this experiment, seven different powder samples were tested, which include: (i) wheat flour, (ii) baking soda, (iii) sucrose crystals, (iv) finely powdered sucrose, (v) salt, and (vi) talcum powder. A seventh trace acts as control data and is obtained with an empty holder. Each powder had prepared samples at three different thicknesses: 2, 3 and 4 mm. The T-ray responses of seven powder samples in the time domain are demonstrated as Fig 9.2(a); these measurements were obtained from the 3 mm samples.

9.2.3 Leave-one-out error estimator

The leave-one-out error estimator is a form of non-parametric error estimation and plays an important role in validation for pattern recognition problems. It evaluates each unknown feature vector and provides a basis to evaluate classifier designs (Fukunaga and Hummels 1989). Under the leave-one-out error estimation procedures, each of the N samples $X_\omega[i], i = 1, ..., N$ from class ω are tested using a nonparametric classifier, in our case, the Mahalanobis distance classifier, trained on the remaining $N-1$ samples $X_\omega[i], i = 1, ..., N-1, \omega \neq i$ (Fukunaga and Hummels 1989).



(a) (b)

Figure 9.2. Measured time domain T-ray signals and phase spectrum. (a) T-ray pulsed signals after transmission through seven different types of powder and the holder. All powders are pressed into 3 mm pellets and are as follows: baking soda, wheat flour, sucrose crystals, finely powdered sucrose, sand, table salt, and talcum powder. (b) Phase plot from the Fourier transform of the de-noised T-ray signal measured from a 3 mm baking soda sample.

9.2.4 Results

Our classification experiments consist of two sets of results, which are chosen to emphasize the comparison of performance with and without wavelet preprocessing. For both cases, two classification methods, labelled by Method (A) and Method (B), are adopted and applied to the T-ray data from seven powder samples at thicknesses of 2 mm, 3 mm and 4 mm. In all cases, the frequency domain amplitude and phase at discrete frequencies are used as features. With preprocessing, the T-ray data are preprocessed using the wavelet techniques detailed in Section 6.5.

Method (A), represented in the Table 9.1, uses the following classification procedure. First, the Mahalanobis distance classifier was trained using the responses from 25 randomly chosen pixels of seven classes of powders at thicknesses of 2 mm and 4 mm. Then, the trained classifier was tested with another 25 randomly chosen pixels from all seven classes of powders, at a sample thickness of 3 mm. The computation procedure here is repeated 50 times to obtain an average accuracy. For Method (B), a leave-one-out (LOO) method was performed 51 times to validate the classifiers. That is, the classification experiment was repeated 51 times, with 50 pixels used for classifier training and one pixel for testing in each case.

In selecting the set of discrete frequencies for use in the experiments, we adopt a search approach. To limit the search scope, we consider a small number of frequencies, and then evaluate different combinations of frequencies. From our search, it is found that only three optimal frequency components are sufficient to achieve very high classification accuracy after the application of de-noising preprocessing. Empirically, it is observed that adding more frequency components to the dimensionality of the feature vectors does not lead to improvements in classification accuracy. The frequency selection process is performed individually for the cases of measurements with and without preprocessing, to ensure comparison between the best possible performances in both cases.

There are four feature sets formed at optimal frequency combinations that are obtained using an iterative algorithm that examines all frequency triplets available. For the raw measurements, the three frequencies are 0.22 THz, 0.45 THz and 1.27 THz; for the preprocessed signals, the three frequencies are 0.19 THz, 0.37 THz and 1.38 THz. In addition to the three frequencies, the classification performance of two-frequency subsets are also examined. As a result, there are four feature sets for each of the cases; these are referred to as combinations 1–4 in the results below.

Experimental results for preprocessed signals

Table 9.1 shows the classification steps and the corresponding classification performance mentioned above by applying an FFT of T-ray signals from the seven powder samples at 2 mm and 4 mm after heuristic SURE soft threshold shrinkage de-noising and deconvolution preprocessing. Two classification methods are compared. The classification accuracy using the leave-one-out error estimator for Method (B) is similar to the results obtained in Method (A).

In order to better understand the current classification algorithm, both the classification performance and the relative visual scatter plots after wavelet filter preprocessing are examined. Two three-dimensional amplitude scatter plots (amplitudes only) in the frequency domain are shown in Fig. 9.3(b). The seven different classes of powder samples are shown to cluster together, with just a small degree of overlap in the 3D plot. The corresponding classification accuracy reaches 98.9%. To examine the effectiveness of the chosen frequencies, the 3 frequency features are further divided into three different subsets of 2 frequencies: amplitudes and phase at 0.19 THz and 0.37 THz, 0.19 THz and 1.38 THz and 0.37 THz and 1.38 THz. The corresponding classification accuracies

Table 9.1. The classification accuracies for the seven types of powder samples. Left resultant column corresponds to the cases with wavelet preprocessing; the right most column corresponds to raw signals. The wavelet preprocessing, here, adopts 'heursure' soft thresholding with a Daubechies 8 (db8) wavelet at 3 DWT levels. The detailed subspaces are applied for the reconstruction of T-rays. Feature extraction identifies amplitude and phase feature subsets at 3 optimal frequencies and at the three different 2 frequency combinations. The four different frequency combinations are labelled as Combinations 1, 2, 3 and 4. The training samples are taken from 2 mm, 4 mm measurements and the test sample has thickness of 3 mm. Two classification methods, Method (A) and Method (B) are demonstrated for the classification accuracy at the four different frequency combinations.

	Experimental parameters			Wavelet preprocessing	Raw signals
Feature extraction	feature			Preprocessed amplitude and phase	Raw amplitude and phase
	method			iterative algorithm	iterative algorithm
	training sample thickness (mm)			2, 4	2, 4
	test sample thickness (mm)			3	3
	Selected frequencies (THz)	comb.	1	0.19, 0.37 and 1.38	0.22, 0.45 and 1.27
			2	0.19 and 0.37	0.22 and 0.45
			3	0.19 and 1.38	0.22 and 1.27
			4	0.37 and 1.38	0.45 and 1.27
Method (A)	training vector dimensions			25 pixels	25 pixels
	test vector dimensions			25 pixels	25 pixels
	classifier			Mahalanobis	Mahalanobis
	number of tests			50 time	50 time
	averaged maximum classification accuracy (%)	comb.	1	98.9	78.6
			2	85.4	51.37
			3	78.4	55.69
			4	74.5	46.11
Method (B)	training vector dimensions			50 pixels	50 pixels
	test vector dimensions			1 pixel	1 pixel
	classifier			Mahalanobis	Mahalanobis
	number of tests			51 times	51 times
	average maximum classification accuracy (%)	comb.	1	98.6	55.18
			2	89.9	52.94
			3	77.0	56.30
			4	77.6	46.5

obtained are 85.4%, 78.4% and 74.5%, respectively. Therefore, the classification accuracy achieved with only two frequencies is at least 15% inferior to the classification accuracy achieved with three frequencies. These results are labelled in Table 9.1 as Method (A). Fig. 9.3(a) projects the 3D scatter plot to a 2D plane with frequency components consisting of 0.19 THz and 0.37 THz. It is clear that Fig. 9.3(b) shows better clustering than Fig. 9.3(a).

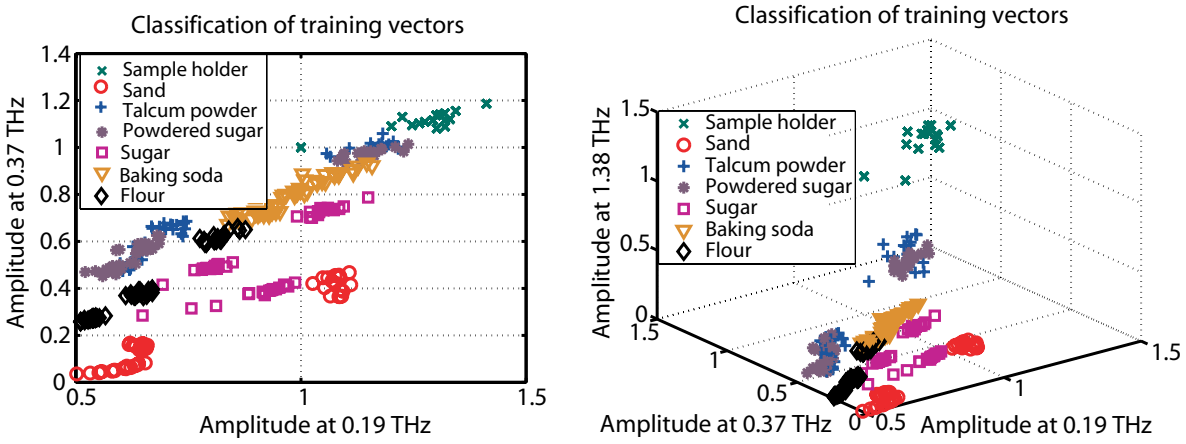
A leave-one-out protocol, labelled by Method (B), is used for comparison to Method (A). Recall that 50 pixels are randomly selected from the 51 pixels of the seven classes of powder data at 2 mm and 4 mm thicknesses, which is used to train the Mahalanobis distance classifier. Then the Mahalanobis distance classifier is tested by the one pixel from the seven powder samples of 3 mm thickness. The procedure is iteratively performed 51 times at each frequency combination—from combinations 1 to 4 (see Table 9.1). The corresponding average maximal classification accuracies are calculated as 98.6%, 89.9%, 77% and 77.6%, respectively.

Experimental results for raw signals

For comparison, the feature sets are extracted for the classification of the raw signals via a similar procedure as described above. A set of three frequencies are selected iteratively, with the amplitude and phase at these frequencies used as features; three 2-frequency subsets are further extracted for classification experiments to justify the need for three frequencies. The averaged maximal accuracy with obtained from amplitudes and phase features at 0.22 THz, 0.45 THz and 1.27 THz is 78.6 %, which is 19 % lower than the averaged accuracy of 98.9% obtained with preprocessed signals. For the other feature sets, the averaged accuracies of the raw signals corresponding to the three 2D features are: 51.37%, 55.69% and 46.11%, respectively, which are inferior to the preprocessed signals by approximately 20% across the board. Compared to the best accuracy achieved with preprocessed signals and 3D features, the 2D raw signal features are lower by 23% at least. The massive discrepancy between the two sets of signals clearly favours the use of wavelet preprocessing when performing classification experiments on T-ray signals.

When a leave-one-out error estimator, labelled by Method (B), is used to validate the classifier obtained for the raw signals, the average classification accuracies are: 55.18%, 52.94%, 56.30% and 46.5%, for the 3D and three 2D feature sets, respectively. There

is obviously a large reduction in classification accuracy compared to the results obtained with wavelet preprocessing. The classification results of the raw signals are summarised in the last column of Table 9.1.



(a) (b)

Figure 9.3. Amplitude scatter plot. (a) Amplitude scatter plots at 2 optimal frequencies: 0.19 THz and 0.37 THz with a classification accuracy reaching 85.4% and 89.9% corresponding to two types of classification methods (a) and (b). (b) Amplitude scatter plots at 3 optimal frequencies: 0.19 THz, 0.37 THz and 1.38 THz with a classification accuracy reaching 98.9%. This figure is plot after 'heursure' soft thresholding with a Daubechies 8 (db8) wavelet at 3 DWT levels.

9.2.5 Conclusions

In this case study, an improvement in classification accuracy is demonstrated by applying wavelet-based techniques to the preprocessing of T-ray pulsed signals to achieve enhanced T-ray classification. The results reveal that a higher level of accuracy can be obtained after implementing heuristic SURE wavelet shrinkage de-noising and deconvolution prior to classification.

9.3 Classification using subspace and wavelet packet algorithms

This work compares classification results of lactose, mandelic acid, and dl-mandelic acid, obtained on the basis of their respective THz responses. The performances of

three different pre-processing algorithms applied to the time-domain signatures obtained using a THz-transient spectrometer are contrasted by evaluating their classifier performances. A range of amplitudes of zero-mean white Gaussian noise are used to artificially degrade the signal-to-noise ratio of the time-domain signatures to generate the data sets that are presented to the classifier for both learning and validation purposes. This gradual degradation of responses by increasing the noise level is equivalent to performing measurements assuming a reduced integration time. Three signal processing algorithms are adopted for the evaluation of the complex insertion loss function of the samples under study: (i) standard evaluation by ratioing the sample with the background spectra, (ii) a subspace identification algorithm, and (iii) a novel wavelet packet identification procedure. Within class and between class dispersion matrices are adopted for the three data sets. A discrimination metric evaluates how well the three classes can be distinguished within the frequency range 0.1 - 1.0 THz using the above algorithms.

The current work proposes the use of a novel system identification scheme implemented in the wavelet domain and contrasts its ability to extract the important features in the signal with that of the N4SID subspace identification method (Ljung 1999). N4SID operates on Hankel matrix equations, illustrated in Eq. (7.18), in order to compute an estimate of the state sequence metric $\mathbf{X}$ from the basic metric equation, $\mathbf{Y} = \mathbf{\Gamma X} + \mathbf{\Omega U} + \mathbf{e}$ of the subspace algorithms. Here, $\mathbf{Y}$, $\mathbf{U}$ and $\mathbf{e}$ are Hankel matrices formed with the output samples, the input samples, and noise, respectively. Note that $\mathbf{\Omega}$ is a Toeplitz matrix. The state sequence metric $\mathbf{X}$ and extended observability metric $\mathbf{\Gamma}$ can be solved via calculating the singular values. The mathematical details are given in Eqs. (7.42)-(7.45).

The goal of this work is to demonstrate efficient and robust classification algorithms that could be adopted by the biomedical and pharmaceutical communities (Zeitler *et al.* 2007b), which are envisaged to provide the technology pull required for the further proliferation of THz-transient spectrometers.

9.3.1 Evaluation of complex insertion loss, subspace and wavelet packet identification

In this experiment we use a commercial T-ray system, the Picometrix T-ray 2000, to perform the THz measurements. The system does *not* use a lock-in technique, because

it is designed to acquire data as fast as possible. This means that the THz data achieved is noisier than if a lock-in technique were used, and this motivates the need for our signal processing techniques described here.

As we used no lock-in technique, there was also no chopper wheel involved in this experiment. A conductive antenna or Auston switch is used in both THz emitter and detector, giving a bandwidth up to 1.7 THz. All time-domain data files are of the same length starting from 50 ps finishing at 605.26 ps. The scan rate (speed) is 5 ps/second. The time interval for each data point is 0.033890 ps.

All samples are pressed to pellets and mounted under the same pressure, with a thickness of 1.67 mm $\pm$ 0.2 mm, in addition to mandelic acid with a thickness of 2.45 mm $\pm$ 0.2 mm. Powders of sample material are mixed with polyethylene (PE) at a ratio of 40:60 and then pressed together. We mix the samples with PE for two reasons. First, it is needed to give the sample some stability, as most of these substances would not be stable when pressed into pellets. The second reason is to dilute the pure sample in order to prevent saturation effects of the peaks.

Time domain responses of lactose, mandelic acid, and DL mandelic acid are recorded using a THz-transient spectrometer, with each response being co-averaged 10 times. Typical background and sample signatures are shown in Fig. 9.4. Linear detrending of the co-averaged experimental data sets using the detrend.m routine in MATLAB is also shown. The aim is to remove the linear trend (the mean value) from the raw measured experimental data.

After the pre-processing procedures, the background and sample responses are employed as input u and output y signals, respectively. The frequency response of an identified model would be an estimate of the complex insertion loss (CIL). The singular value plot generated in the subspace identification procedure is presented in Fig. 9.5. Following the default recommendation of the N4SID function, a 4^{th} order model is adopted. It is worth noting that when using MATLAB's function n4sid.m, different results are obtained by pre-establishing an order of 4 or by making such a choice after testing orders 1 to 20. The results presented in this work are obtained by testing orders 1 to 20. The resulting CIL, which corresponds to the frequency response of the 4^{th} order model, is presented in Fig. 9.5 (b) in the spectral range 0.1 - 1.0 THz. The result obtained by ratioing the sample spectrum against the background spectrum is also shown. This processing is realised via using an asymmetric (Mertz) triangular apodization window for the FFT calculations. As can be seen, by using the subspace

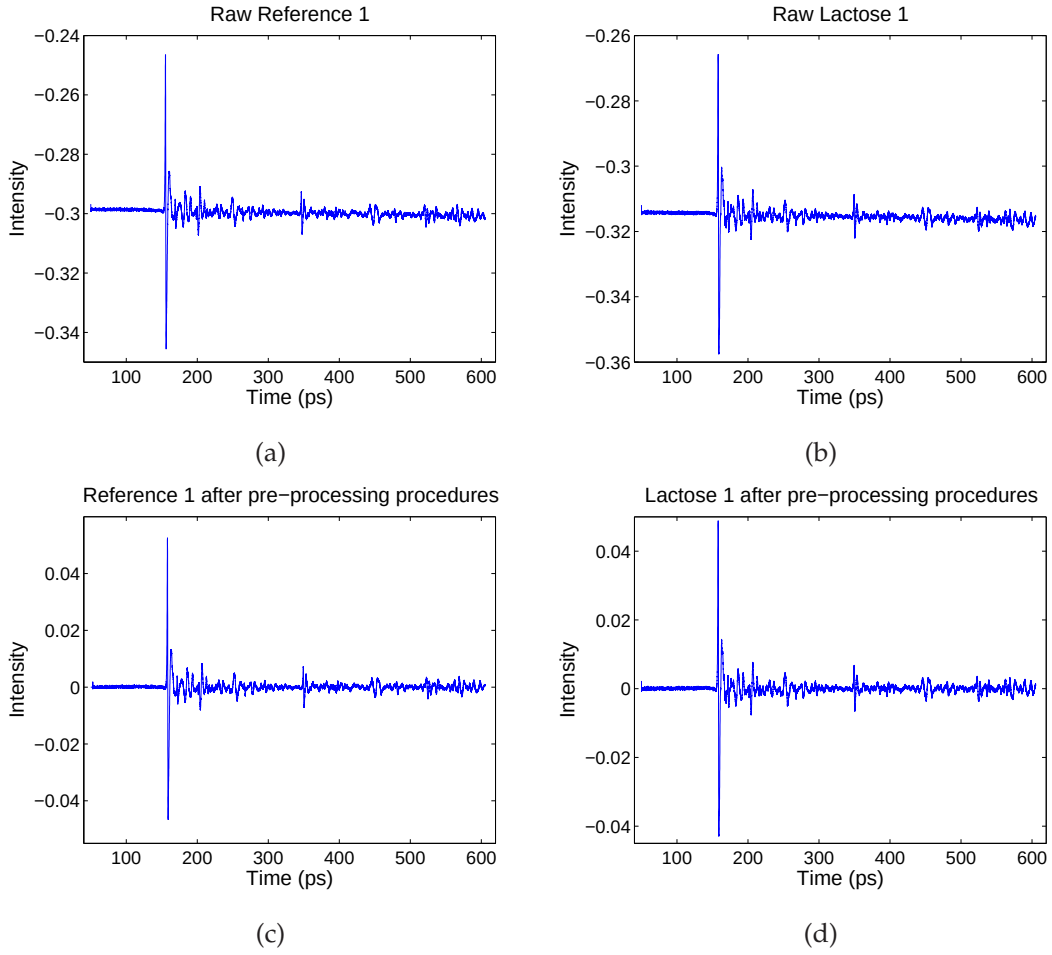
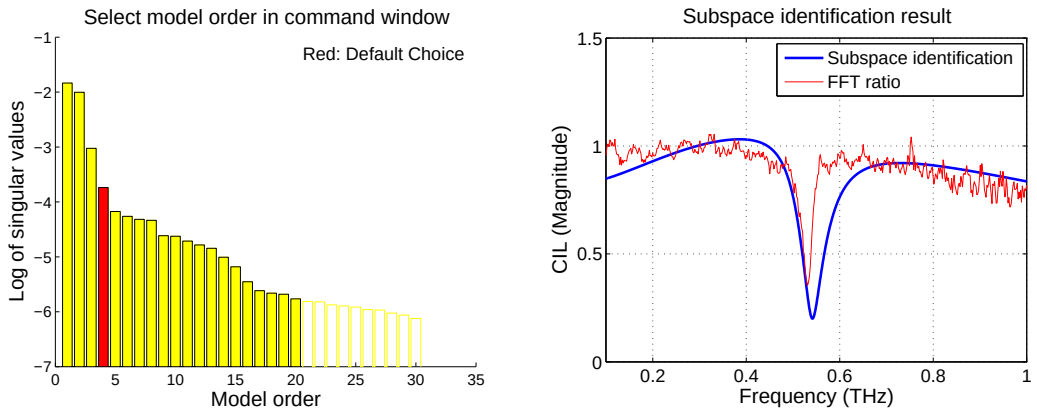


Figure 9.4. THz measurements of a lactose sample. Background (reference) and sample (lactose) responses before and after the pre-processing procedures (detrending and alignment).

algorithm, the estimated position of the absorption band is slightly biased towards higher frequencies, and its magnitude seems to be over-estimated.

Prior to the wavelet packet identification procedure, a 6^{th} order Butterworth band-pass filter is employed to band-limit the responses to the 0.1 THz - 1.0 THz range. The responses are then re-sampled in order to reduce the sampling frequency by a factor of 8. This procedure is used to reduce the number of wavelet decomposition levels required to attain an appropriate frequency resolution. The settings for the wavelet packet identification procedure use a db12 wavelet with a maximum tree depth of 9 decomposition levels (including the root node); values tested for the s parameter (exponent of the integrator term in the sub-band models) are -1, 0, +1. Fig. 9.6(a) presents



(a) (b)

Figure 9.5. Subspace identification of lactose responses. (a) Singular value plot for the subspace identification and (b) calculated magnitude of the complex insertion loss (CIL) for the lactose sample as a function of frequency obtained by ratioing the sample spectrum against the background spectrum (red line) and by subspace identification (blue line).

the resulting wavelet packet tree obtained by the generalized cross-validation procedure. The tree is deeper in a particular frequency range, which actually corresponds to the absorption valley, as shown in Fig. 9.6(b)—the segmentation is more refined in the frequency region corresponding to deeper levels of the tree. It is worth noting that the tree structure is automatically defined by the identification algorithm, with no prior knowledge of the spectral features of the sample under consideration.

9.3.2 Signal processing assuming noisy background and sample response

The standard deviation of the signal noise (white, zero-mean Gaussian) is varied in software from 10^{-4} to 10^{-3} to evaluate the discrimination metric (described in Sec. 8.2) for different signal-to-noise ratios. Ten noisy sample/background signals pairs are generated for each species (lactose, mandelic acid, DL mandelic acid). Therefore, an overall set of 30 complex insertion loss (CIL) functions are calculated for each noise level and for each processing technique (FFT, subspace, wavelet packet). Each of these calculated CIL functions are termed an 'object'. As part of the pre-processing procedure, the time-domain responses are aligned with respect to each other. Fig. 9.7 compares the noisy responses (noise standard deviation of 10^{-3}) before and after the

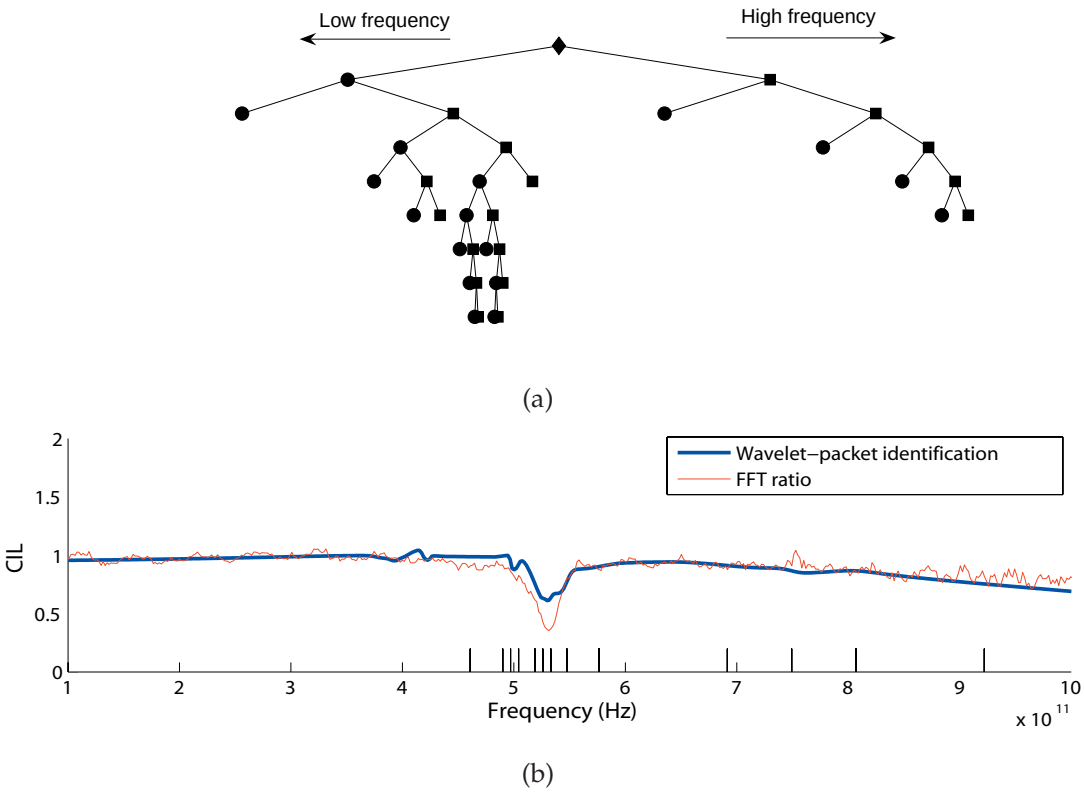


Figure 9.6. Wavelet-packet identification of lactose responses. (a) Resulting wavelet packet tree and (b) CIL for a lactose sample obtained by wavelet packet identification (blue line). The FFT ratio result (red line) is also presented for comparison. The frequency-domain segmentation automatically defined in the identification procedure is indicated by vertical lines at the bottom of the graph. As can be seen, the segmentation is more refined in the spectral region corresponding to the absorption band.

pre-processing procedures. Furthermore, an asymmetric (Mertz) triangular apodization window is used for the FFT calculations.

The singular value plot generated in the subspace identification procedure is presented in Fig. 9.8(a). A 4^{th} order model is adopted, as required by the N4SID function. The resulting CIL, which corresponds to the frequency response of the 4^{th} order model is presented in Fig. 9.8(b). As can be seen, the identification result is very sensitive to the additional noise present in the time domain signatures.

Fig. 9.9 presents the resulting wavelet packet tree obtained after the inclusion of artificial noise. As can be seen, the tree has much fewer nodes as compared to the tree

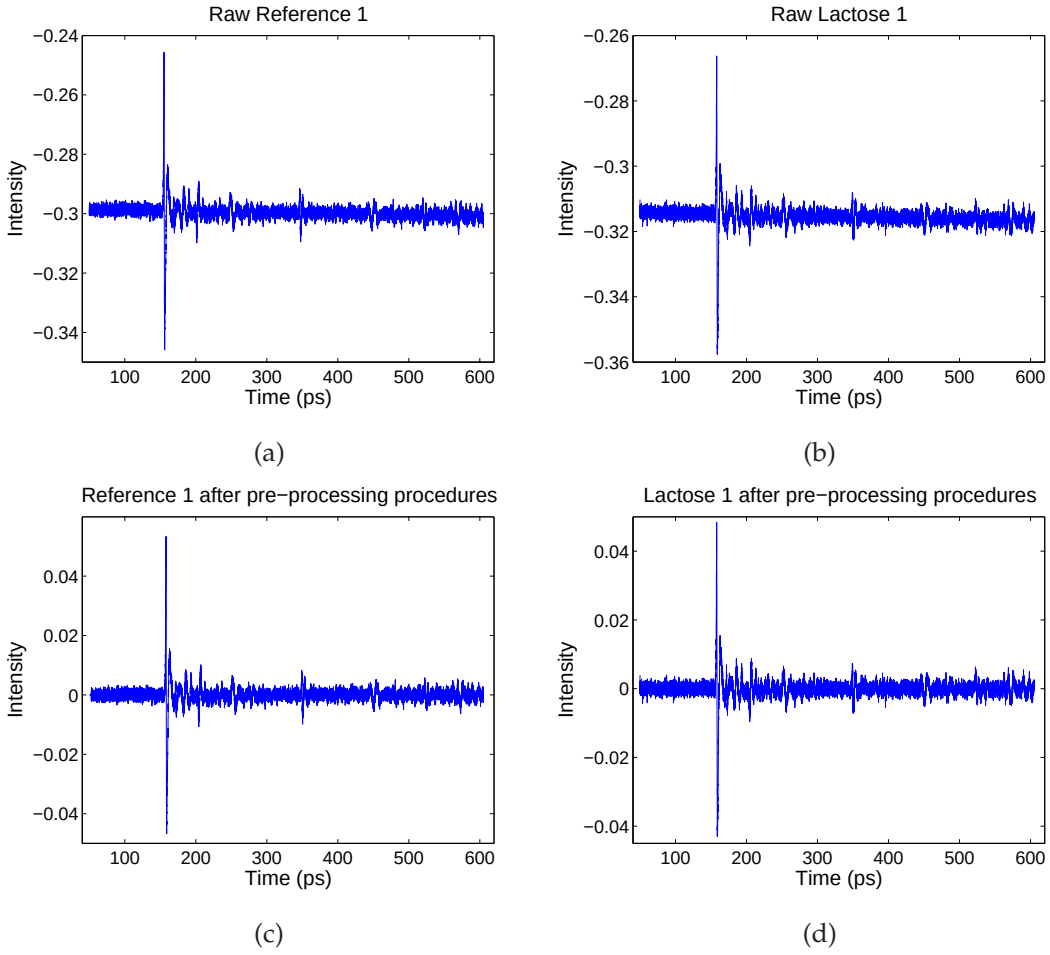
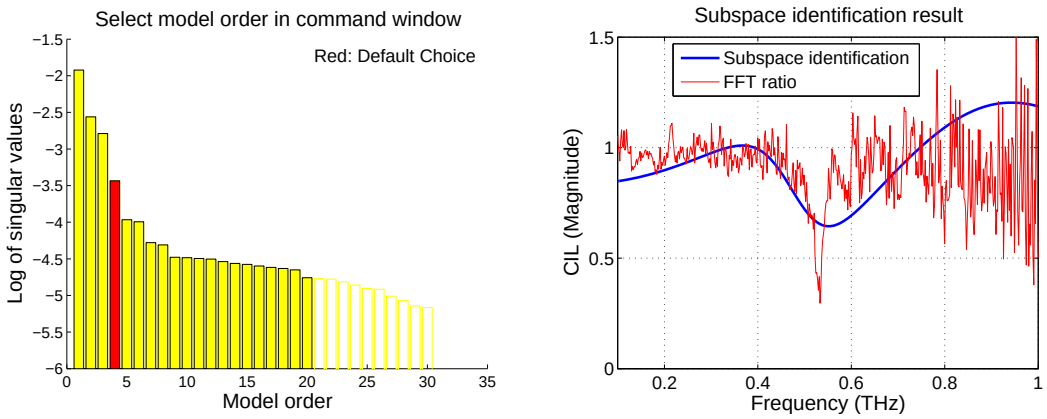


Figure 9.7. Degradation of lactose responses. Noisy background (a) and (c) and sample responses (b) and (d) before and after the pre-processing procedures, respectively.

obtained in the previous case (Fig. 9.6). Such a result is obtained because the generalized cross-validation procedures tend to generate more parsimonious models (i.e. tends to group frequency segments together in the identification procedure) when the signal-to-noise ratio is worse. Again, it is worth noting that the segmentation in the frequency domain is established in an automatic manner, and no prior knowledge of the signal-to-noise ratio is required. This result is more clearly demonstrated in Fig. 9.10, which presents trees obtained for different realizations of noise with standard deviations of 10^{-3} and 10^{-4}. The structure of nodes corresponding to the absorption feature is always present, but the increase in the noise level leads to the pruning of other parts of the tree.



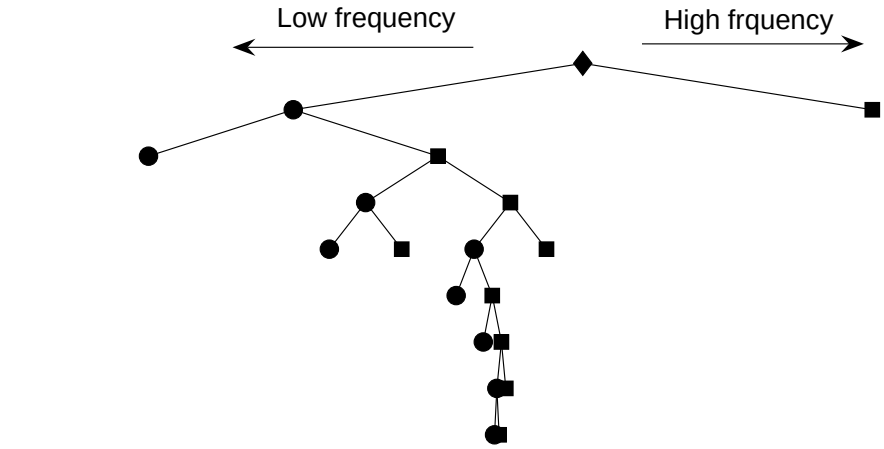
(a) (b)

Figure 9.8. Subspace identification of noisy lactose responses. (a) Singular value plot for the subspace identification after the inclusion of artificial noise and (b) the corresponding calculated magnitude of the complex insertion loss for the lactose sample as a function of frequency obtained by ratioing the sample spectrum against the background spectrum and by subspace identification (4^{th} order model).

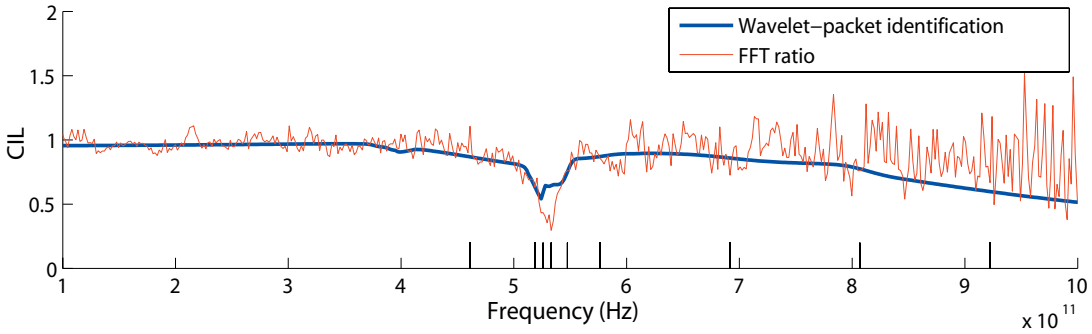
9.3.3 Evaluation via the discrimination metric

In what follows, classes 1, 2, and 3 will refer to the objects corresponding to lactose, mandelic acid, and dl-mandelic acid, respectively. For each noise level and for each processing technique (FFT, subspace, wavelet packet), a discrimination metric is calculated on the basis of the estimated CIL magnitude in the range 0.1 - 1.0 THz. To do so, $x_{m,n}$ is viewed as the CIL magnitude of the m^{th} object $(m = 1, \cdots, 30)$ at the n^{th} spectral bin, and assume 500 spectral bins uniformly distributed in the range 0.1 - 1.0 THz (that is, $n = 1, \cdots, 500$).

The discrimination metric is a key issue in the current machine learning algorithm. This work considers the general problem of learning from pairwise constraints in the form of must-links and cannot-links. As one kind of side information, a must-link indicates the pair of the two data points must be in a same class, while a cannot-link indicates that the two data points must be in two different classes. Given must-link and cannot-link information, our goal is to learn the Euclidean discrimination metric. Under this metric, we desire the metric value of point pairs in must-links are as large as possible and those of point pairs in cannot-links are as small as possible. The larger value of resultant metric $\mathbf{F}$, corresponding to these must-link point pairs, means that



(a)



(b)

Figure 9.9. Wavelet-packet identification of noisy lactose responses. (a) Resulting wavelet packet tree after the inclusion of artificial noise and (b) CIL for the lactose sample obtained by wavelet packet identification after the inclusion of artificial noise (blue line). The CIL result calculated using the ratio of sample and background FFTs is also presented for comparison (red line). The frequency-domain segmentation automatically defined in the identification procedure is indicated by vertical lines at the bottom of the graph. As can be seen, fewer frequency segments were employed, compared to the results in Fig. 9.6.

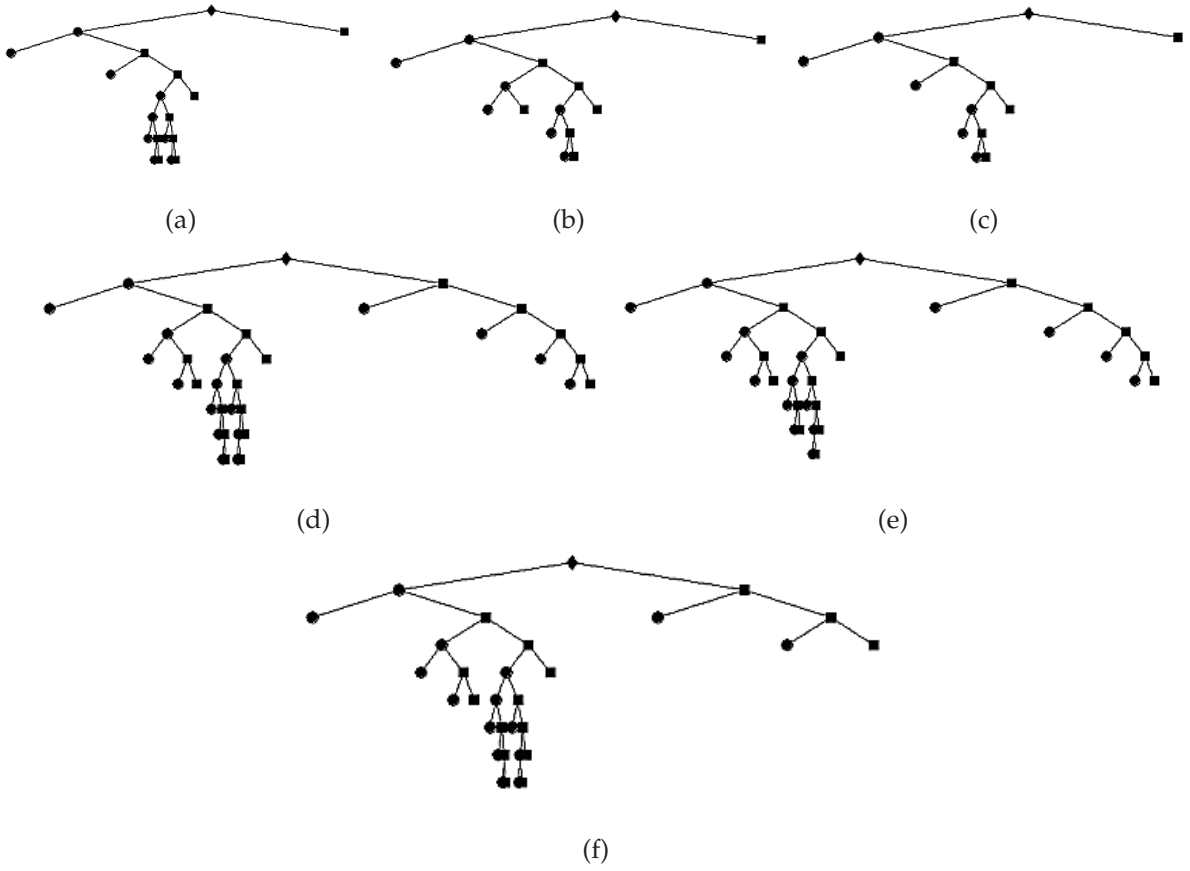


Figure 9.10. Wavelet-packet identification of lactose responses with different noise levels. Trees obtained with different realizations of noise with a standard deviation of 10^{-3} ((a), (b), (c)) and 10^{-4} ((d), (e), (f)).

these must-link point pairs can be well discriminated from those cannot-link ones, and *vice versa*.

In order to observe the performance of discrimination metric with a variety of noise levels, we plot resultant discrimination metric values versus different noise levels superimposed on the THz-pulse data sets. According to this metric, the identification methods are seen to be more robust with respect to noise than the standard ratioing procedure. In particular, the proposed wavelet packet identification technique becomes slightly superior to the subspace method at larger noise levels. But for small noise levels, the advantages of using the identification algorithm are unclear.

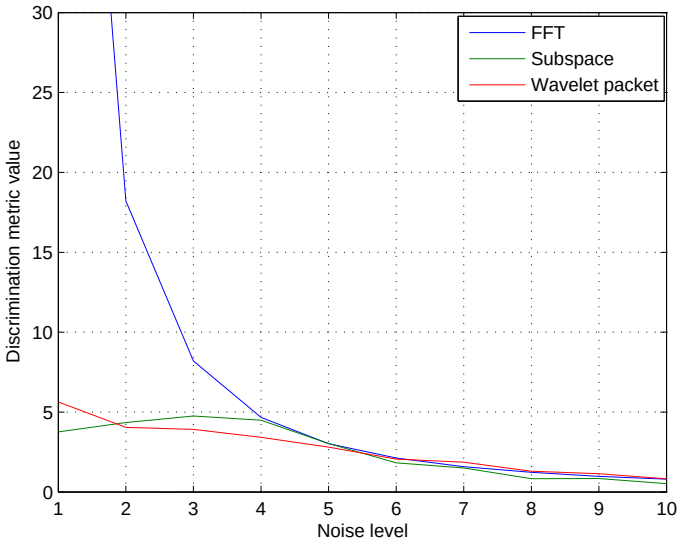


Figure 9.11. Plots of resultant discrimination metric. Plots of the adopted discrimination metric F for different noise levels ($\times\ 10^{-4}$) superimposed in the THz-transient data sets.

9.3.4 Conclusion

A wavelet packet identification scheme for discriminating between lactose, mandelic acid, and dl-mandelic acid THz transients is proposed. After the subsequent evaluation of the complex insertion loss using the output of the model as opposed to the direct ratioing of the spectra, we observe that a ratio composed of the model output smooths out the calculated value of the complex insertion loss function across the frequencies range of interest and identifies more correctly the absorption bands of the samples than the subspace algorithm. This is the case even when the time-domain signatures are corrupted by additional noise. Within class and between class dispersion discrimination matrices are adopted to evaluate the benefits of the proposed algorithm in classification tasks. The results are more robust with respect to noise than those obtained by the standard ratioing procedure, but the advantages of using the identification algorithm for small noise levels are unclear. It is possible that greater benefits could be obtained for more spectrally rich samples, as the wavelet packet technique has been shown to be particularly suited to the identification of systems with several spectral resonance features (Paiva and Galvão 2006).

9.4 Application of AR models of wavelet sub-bands for classification

Recent advances in T-ray sources and detectors have made it possible to image opaque objects and discriminate tumour cells from normal tissue (Löffler *et al.* 2002). While much effort has been devoted to improving the signal to noise ratio and repeatability of measurements as well as reliability in the function of the spectrometers, the further processing of THz transients has only recently received some attention in the literature (Löffler *et al.* 2002, Yin *et al.* 2006). T-ray classification relies in observing changes in pulse amplitude, phase as well as dispersion characteristics of the sample under study. The pulsed nature of time-domain signals obtained in T-ray spectrometry justifies their decomposition in the wavelet domain as it can provide better de-noising capabilities (Ferguson and Abbott 2001a). Furthermore, compared to Fourier-based techniques, a wavelet decomposition of the experimental signal can provide better time-frequency localization characteristics (Donoho 1995), which better facilitates the subsequent classification task.

This Section introduces a novel parametric modelling procedure to the wavelet decomposed T-ray pulsed signals that improves feature extraction and classification. The success of the proposed algorithm in classifying human bone osteoblasts (HBO) against human osteosarcoma cells (HOS) as well as at differentiating between six types of powder samples is presented as a case study. The choice of biological samples is made on the basis that there have been suggestions that THz transient spectroscopy can be used for the early detection of cancerous tissue (the contrasting mechanism being the enhanced blood circulation in the cancerous tissue). The reported work complements previous work in classifying basal cell carcinomas, (a form of skin cancer), which has been conducted by researchers at the University of Cambridge and TeraView Limited (Woodward *et al.* 2002). It needs to be emphasised that the rationale for picking bone cancer cells is that techniques for culturing bone cells on a Petri dish are well-established and readily accessible. Our positive results with T-ray detection of cancerous bone cells will motivate future research to explore other classes of cancer cells.

We also present a case study using powder samples. The motivation for using THz pulses for extracting information on densities, thicknesses, and number of absorber

molecules per unit volume in different powder samples stems from the fact that substance detection is an increasingly important area in the pharmaceutical—drug polymorphs and isomorphs (Strachan *et al.* 2005, Zeitler *et al.* 2007b, Watanabe *et al.* 2003)—as well as security industries, e.g. spectral fingerprinting of explosives and illicit drug detection (Federici *et al.* 2005, Kawase *et al.* 2003). Our goal is to demonstrate a feature extraction methodology that is non-specific to the data sets in hand. This is of significant importance to the THz community as data driven classifiers prohibit proper inter-comparison between results obtained in different labs and therefore preclude the development of standards, guidelines and specifications that could be adopted by the biomedical, pharmaceutical as well as security sectors, which are envisaged to become emerging markets for THz-pulsed spectrometers.

9.4.1 Terahertz pulse measurements

A further advantage in using THz pulses instead of infrared as a measurement modality is the fact that the wavelengths are longer, and noise due to the motion of the translation stage is sufficiently small compared to the wavelength permitting the extraction of phase information. A limitation of current measurement techniques is the low power per spectral bin and consequently, the small signal to noise ratio in the measured complex insertion loss of the sample. It has previously been shown that a combination of wavelet transform techniques and statistical models can mitigate the effects of noise and extract effective features (frequency dependent dispersion, attenuation and phase delay properties of the sample) for classification (Ferguson and Abbott 2001a). Normal human bone (NHB) osteoblast cells were obtained from patients and cultured from small pieces of trabecular bone for 4-6 weeks to obtain a confluent culture. Human osterosarcoma (HOS) cells were cultured from an immortalised cell line. The confluent culture was obtained within one week. A 25 ml polystyrene flask with flat bottom was used to culture the cells under a 5% carbon dioxide environment and a temperature of 37°C. In order to perform T-ray imaging, the rectangular flasks were tipped and placed in the T-ray (x-y) translation stage. A T-ray image was obtained at 10 different positions to provide spectroscopic data, with a distance interval of 50 μm. The above procedure was performed for each of the three flasks—normal cells, cancerous cells and the container with media solution—and iterated a further 5 times until 50 pixels of T-ray responses were obtained for all three flasks (Ferguson *et al.* 2004). This was

regarded as a sufficient amount of data for a verification of classification effectiveness for different cells.

For the powder sample classification experiment, six different powdered substances are used: sand, talcum, salt, powdered sugar, wheat flour, and baking soda with thicknesses of 2 mm, 3 mm and 4 mm. The variations in thickness enabled investigation into thickness-independent classification. A seventh set of measurements of an empty sample holder was used as a reference to provide a background spectrum. The sample holder consisted of two Teflon blocks separated by a translation-stage based control mechanism. This guaranteed a consistent powder density. The Teflon sample holder was mounted on an x-y translation stage. Since Teflon is dispersionless and has a very low absorption coefficient at THz frequencies, there is minimal distortion of the T-ray pulse as it propagates through the holder. A 2D T-ray image of the sample was obtained via a raster scan; such an image allowed the effects of different scattering paths and minor variations in powder thickness and density to be observed. The data in this work was obtained from a truncated raster scan with measurements taken from 50 co-linear locations, or pixels (with a spacing of 100 μm). The integration time of 30 minutes per sample corresponded to an integration time of 36 s per pixel.

9.4.2 Motivation

Fourier transformation of the TPI data described in the previous section provides the frequency dependent characteristics of a target sample. The sampled T-ray pulses are a function of discrete time, to which the discrete wavelet transform (DWT) is applied. To realise the DWT, simple digital filter banks are utilised in a recursive structure to calculate wavelet transform coefficients of T-ray signals. For convenience we recall a number of points already made in Chapter 6 and build on these. Formally, the representation of signals with their wavelet transform coefficients is known as a multi resolution analysis (MRA). The theory undelying MRA allows a systematic method for constructing (bi)orthogonal wavelets (Daubechies 1988) and leads to the fast wavelet transform (FWT), also known as Mallat's algorithm (Qian 2002, Mallat 1989). In principle, wavelet-based techniques are very well suited to studies of non-stationary time-domain data sets, highlighting the variability of features at different time-frequency scales (Mallat 1999, Donoho 1995). The main concern about the current T-ray TPI measurements is its corruption by different types of noise which limit the practical usefulness of this mode of imaging. One of the main problems is the distortion of the

T-ray pulse as it propagates through the optical system. As a result, the received T-ray signals are strongly dependent on the acquisition conditions, and similar samples may produce variations in the measured signals in different portions of the image (Ferguson and Abbott 2001a). In order to keep image acquisition within realistic timescales, a short integration time per pixel is adopted as common practice, this motivates the need for the THz-transient de-noising process. The use of perfect reconstruction quadrature mirror filter banks has been extensively discussed by Vaidyanathan (1993) for the purpose of de-noising and generating bases of compact support. The works of Vetterli and Kovacevic (1995) as well as that of Strang and Nguyen (1996) further complement the above, elaborating more on sub-band coding. Sherlock and Monro (1998) discuss how to apply Finite Impulse Response (FIR) filters of arbitrary length to describe the space of orthonormal wavelets, further parameterizing the wavelet coefficients at each decomposed level. Tuqun and Vaidyanathan (2000) propos a state-space approach to the design of globally optimal FIR energy compaction filters. Since, in our work, there is no requirement for adopting an algorithm with a perfect reconstruction property, as our ultimate goal is feature extraction and classification, our constraints are more relaxed compared to those used in filtering or signal compression applications. Divine and Godtliebsen (2007) suggest that for feature exploration purposes, it is possible to assume stationarity over some time interval and smooth the wavelet spectrum along the time axis using an Auto Regressive (AR) model. Paiva and Galvão (2006) also discuss a wavelet packet decomposition tree algorithm that establishes frequency bands where sub-band models are created. Both approaches propose the modelling of the approximation and detail wavelet coefficients in order to further extract statistically significant features and a similar approach is adopted in our work.

A typical de-noising procedure consists of decomposing the original signal using the discrete wavelet packet transform (DWPT) or the discrete wavelet transform (DWT) (Mallat 1999, Daubechies 1992, Jensen and la Cour-Harbo 2001, Percival and Walden 2000), thresholding the detail coefficients, and reconstructing the signal by applying the appropriate inverse discrete wavelet transform (IDWT) or inverse discrete wavelet packet transform (IDWPT). In our work we adopt (i) the bior6.8 (DWPT) and (ii) the db20 (DWT) wavelet families for de-noising after comparing the classification results obtained using the following: db1, db8, db20, sym1, sym2, sym4, sym8, sym12, coif2, coif5, bior1.1, bior2.8 and bior6.8 and adopting a three level decomposition. These

wavelets are chosen as representatives for general classes of non-orthonormal, orthonormal, and biorthogonal wavelets, respectively. For the de-noising of femtosecond THz transients, a three-level decomposition is usually sufficient (Hadjiloucas *et al.* 2004) and unnecessary computational load associated with more decomposition levels can be avoided. The Stein's Unbiased Risk Estimate (SURE) and the 'heuristic' SURE methods (Donoho 1995) are used separately to estimate the soft threshold parameter (λ^S) for the cancer and powder classification experiments, respectively.

9.4.3 Resultant THz experiments

Wavelet preprocessing

Fig. 9.12 illustrates the performance of wavelet de-noising, after adopting the WP SURE denoising procedure for both normal (NHB) and cancerous (HOS) cells and performing de-convolution to eliminate the features due to the container. The background reference signal is used for the de-convolution processing, which isolates the TPI sample response from that of the container. The de-convolution procedure is performed by dividing the respective Fourier transforms of the sample signals with that of the background time domain signal at the corresponding pixel.

The comparisons of the signals in the time domain are performed for one pixel of normal and cancerous cell responses together with their containers before and after employing WP SURE denoising. The effectiveness of wavelet package SURE soft threshold shrinkage de-noising is demonstrated in the three visually separable T-ray pulsed responses of Fig. 9.12(b), in comparison with the original measured T-ray signals of Fig. 9.12(a).

Classification of NHB and HOS cells

For the classification of cancer cells, the averages of two different orders of AR coefficients are extracted for use as features. Wavelet transform depths of 2 and 3 are compared, and the biorthogonal spline wavelet of order 6.8 is used to compute the discrete wavelet packet transform. A Mahalanobis distance classifier is trained using a

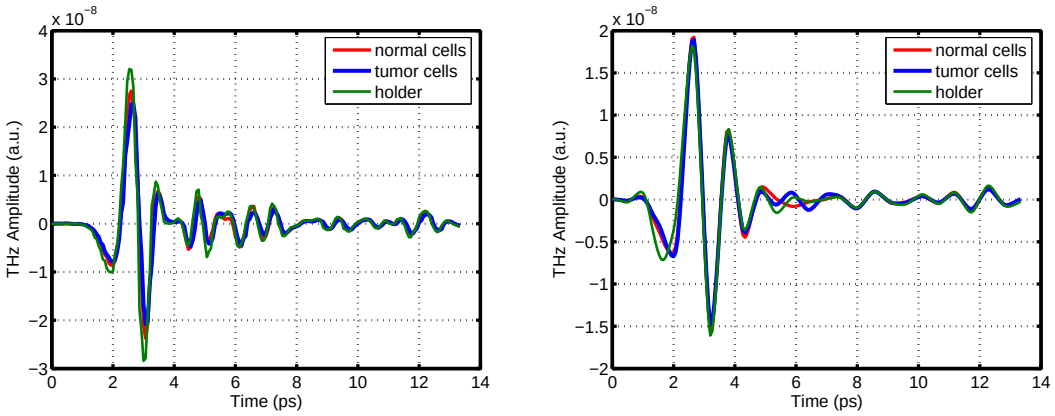


Figure 9.12. Comparisons of the time domain signals for NHB and HOS cells. The plots are for a single pixel of T-ray images for a normal cell (NHB) and a cancerous cell (HOS) together with the empty container (holder) (a) before and (b) after employing WP SURE denoising of bior6.8 wavelet.

portion of the T-ray responses corresponding to each of the three classes—cancer cells, normal cells and the reference signals—then the remainder of the pixel responses are classified. The resultant classification accuracy is used to measure performance. In addition, error prediction covariance at different orders of AR modelling and the different levels of wavelet transform are calculated and employed as another matrix for comparison.

Fig. 9.13 is a block diagram specific to the algorithm for feature extraction and classicization. This current algorithm is also validated again via the classification experiment of powder specimens.

Table 9.2 shows squared error variances of seven orders of the AR model corresponding to the approximation coefficients generated using a bior6.8 wavelet family with 3 levels of wavelet decompostion on an arbitrarily chosen T-ray response. The columns of this table correspond to the AR model order, while the rows correspond to the levels of DWPT prior to calculating the AR models. Data are scaled by a factor of 10^{-16}, demonstrating a very successful modelling. It can be seen that the value of the error variance decreases with increasing AR order. Increasing the DWPT depth tends to increase the value of the error variance. The absolute error variance spans over two orders of magnitude, demonstrating that the choice of the number of decomposition

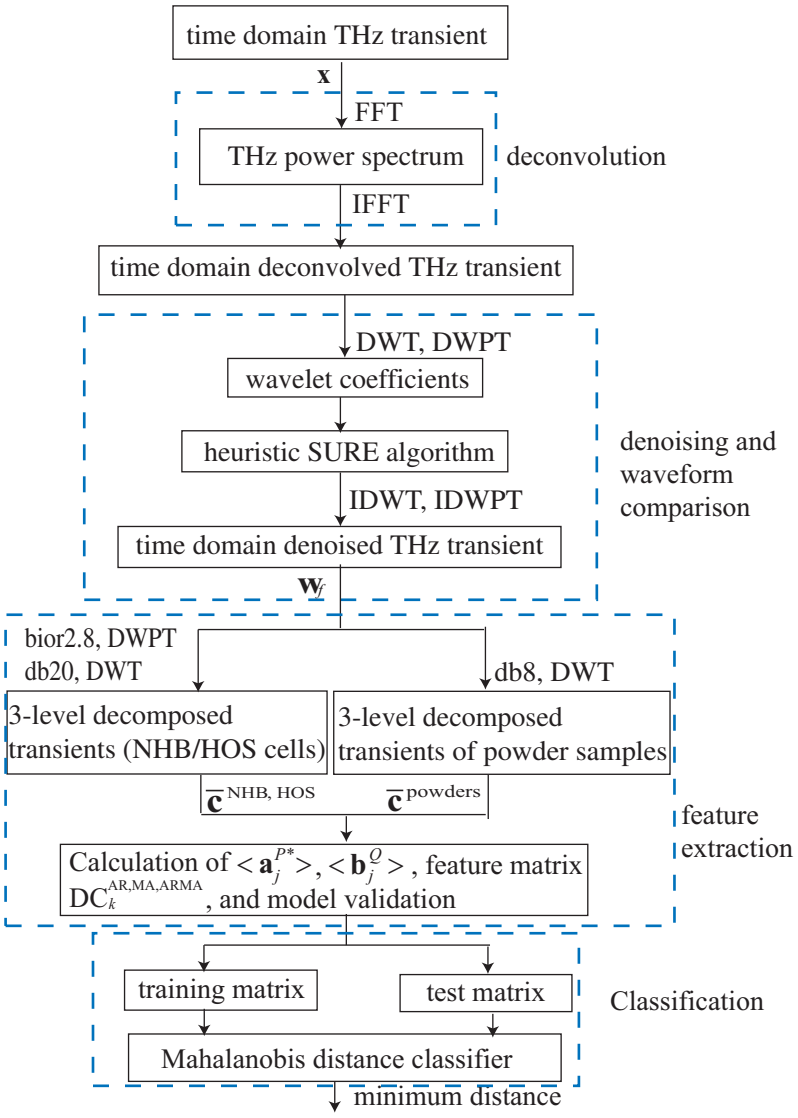


Figure 9.13. Illustration of the current classification algorithm. AR/ARMA modelling and wavelet decomposition are applied on THz measurements.

levels and the AR model order adopted can have a significant effect on the validation process of the model.

Table 9.3 shows the classification accuracy results obtained from two classification methods, which are labelled Method (A) and Method (B) for notational simplicity. In Method (A), the classifier is trained using half of the pixel responses (chosen in an arbitrary manner) for the 3 data classes (normal, cancerous cells, and reference TPI responses), and tested by the remaining half of the pixel responses. In Method (B), a

Table 9.2. The calculated squared error variances on cell samples. Squared error variances of AR models are fitted to DWPT subbands on cell samples, (results scale by a factor of 10^{-16}).

DWPT depth	AR model order						
	2	3	4	5	6	7	8
1	0.01155	0.00767	0.00759	0.00722	0.00426	0.00318	0.00316
2	0.1596	0.1403	0.0763	0.0685	0.0512	0.0480	0.0435
3	0.2202	0.2032	0.2013	0.2011	0.2010	0.2010	0.2010

leave-one-out (LOO) error estimator is used. The leave-one-out estimator uses the Mahalanobis distance classifier, which is trained using a randomly selected set of $N-1$ responses, out of a total of N responses for the three classes and tested using the remaining response. The procedure is repeated N times to obtain the average classification accuracy. In the experiments, all pairwise combinations of AR orders are considered, which implies a total of $\begin{pmatrix} 7 \\ 6 \end{pmatrix} = 42$ combinations. From Table 9.3, it can be seen that the classification accuracies obtained are in the 79.2% to 93.1% range for Method (A) and 71.2% to 90.6% range for Method (B). Considering the results in Table 9.2, Table 9.3 and Table 9.4 together, it can be observed that there is a trade off between the classification accuracy and model complexity. Combinations of lower-order AR models generally outperformed combinations with high-order models. The best classification accuracy is achieved by performing a combination of the features generated from AR models of orders 3 and 5. The classification result after adopting the LOO method reaches 90.85%.

Classification of powder samples

In the powder classification experiment, the responses are obtained from samples of different thicknesses: 2 mm, 3 mm and 4 mm. All the data from the six-classes of powders plus the background reference signal at thicknesses of 2 mm and 4 mm are used to train the classifier— Mahalanobis distance classifier, and all the data of powder samples at a thickness of 3 mm are used to test the classifier. After wavelet denoising using the SURE procedure, a db8 discrete wavelet transform is applied to the denoised powder data. Subsequently, ARMA modelling is performed to extract the relative feature matrix. The performance of Durbin's algorithm for ARMA modelling is characterised by the squared error variances. These results are presented in Table 9.4. The correlation

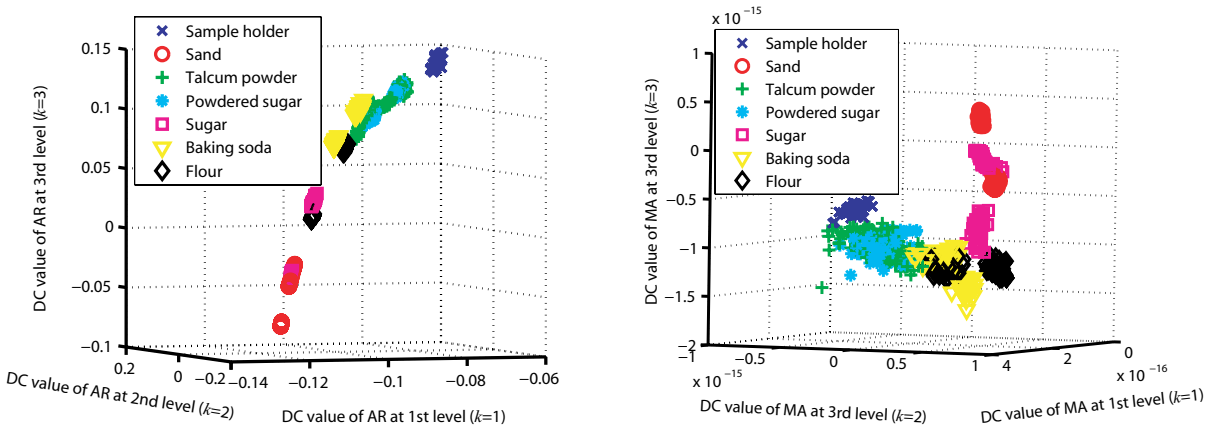
Table 9.3. Percentage classification accuracy of T-ray pulses travelling through cell samples. The range of AR model orders considered is 2 to 8; 3 levels of DWPT with a bior6.8 wavelet are used throughout. Above the diagonal of this table are results from Method (A); below the diagonal are results from Method (B). The bolded values show the best and worst performance.

AR order	AR order						
	2	3	4	5	6	7	8
2		88.9	91.7	90.3	91.7	87.5	87.5
3	**90.6**		90.3	**93.1**	86.1	87.5	88.9
4	87.7	**90.6**		83.3	88.9	84.7	87.5
5	87.0	**90.6**	84.1		86.1	80.6	87.5
6	87.7	**90.6**	87.7	84.8		86.1	84.7
7	87.7	86.2	86.2	83.3	82.6		**79.2**
8	87.0	85.5	84.8	84.1	86.2	**71.7**	

method allows the lowest error variance among the methods investigated, shown in Table 9.5. The error variance of Durbin's method is much lower than the basic Prony's method, though a little higher than that of the correlation method. However, the low variance achieved by the correlation method comes at a cost of a larger model order (AR only).

The ARMA model produces features that separate the classes quite effectively. The classification accuracy in Table 9.8 improves by at least 3% over either AR or MA models alone, as illustrated in Table 9.6 and Table 9.7, respectively. As mentioned in Chapter 6 (Sec. 7.4.2), Durbin's algorithm is used to estimate MA parameters, which is an improvement of the basic Prony method. Compared to the correlation and normal Prony's methods used to fit an AR or ARMA model, that produces a maximum classification accuracy of 96 % and 95%, respectively, Durbin's algorithm improves the maximum classification accuracy by 2%.

Fig. 9.14 shows the scatter plots of the feature vectors for 700 random samples at thicknesses of 2 mm and 4 mm. In order to better understand the learning vectors and show a better grouped data set for each powder sample, we choose the second and fifth order ARMA model at three levels of wavelet decomposition to generate the scatter plots. The seven classes (comprising of the six powder samples and a reference) in Fig. 9.14(a), which uses AR model coefficients, are grouped together, although the class populations show some degree of overlap. The scatter plot in Fig. 9.14(b), which uses MA model coefficients, is not clearer than the case with the AR model coefficients. However, it is found that the combination of the two model coefficients to generate an



(a) (b)

Figure 9.14. Scatter plots of learning vectors. Scatter plots show the learning vectors of discriminating powders with the thickness of 2 mm and 4 mm, corresponding to (a) second order AR model; and (b) fifth order of MA model, over three steps of wavelet transforms.

Table 9.4. The calculated squared error variances on powder samples. Squared error variances of ARMA are fitted to DWT subbands on T-ray signals of powder samples (scaled by a factor of 10^{-17}).

DWPT depth	ARMA model order						
	2	3	4	5	6	7	8
1	0.2405	0.1397	0.1337	0.1147	0.0948	0.0935	0.0797
2	4.458	2.854	1.280	0.976	0.674	0.673	0.660
3	32.09	32.06	26.97	26.93	24.35	24.06	23.06

ARMA feature matrix at the two specified orders actually led to good classification performance, with classification accuracy of 96%.

The ARMA classification experiments are performed on a computer equipped with a 2.4 GHz Pentium4 CPU. The average time spent classifying the six-classes powdered samples and their reference is 51.2 seconds for 42 runs, or 1.2 s per run. It is therefore feasible to use this ARMA modelling algorithm to perform near real-time classification of pulses.

Table 9.5. Squared error variances of ARMA on powder samples. Squared error variances of ARMA are obtained from the correlation method, Prony's method, and Durbin's method, as applied to powder samples.

Method	Correlation	Prony's	Durbin's
Minimum value	0.3399×10^{-19}	0.3581×10^{-4}	0.797×10^{-17}
Maximum value	0.1153×10^{-16}	0.0115	0.3209×10^{-15}

Table 9.6. Percentage classification accuracy of several powder samples. The percentages of accuracies are obtained using an AR modelling matrix.

AR order	AR order					
	3	4	5	6	7	8
2	88.8	75.9	96.1	92.7	95.0	92.7
3		85.4	89.1	88.8	81.8	96.9
4			88.2	96.6	85.2	96.9
5				81.2	75.6	86.0
6					75.9	91.3
7						75.63

9.4.4 Conclusion and future work

The use of an orthogonal transform such as the wavelet transform is well justified for decomposing the time-domain signals obtained using a THz pulsed spectrometer. The further modelling of wavelet coefficients to produce new feature vectors has shown to produce new feature vectors rich in information content irrespective to the morphology of the samples. The use of Durbin's algorithm in estimating the ARMA

Table 9.7. Percentage classification accuracy of several powder samples. The percentages of accuracies are obtained using an MA modelling matrix.

MA order	MA order					
	3	4	5	6	7	8
2	84.9	84.9	81.5	84.9	85.2	84.0
3		85.7	78.7	72.0	77.0	79.8
4			73.4	79.8	80.1	81.0
5				72.0	81.0	79.3
6					70.6	77.9
7						73.1

Table 9.8. Percentage Classification accuracy of several powder samples. The percentages of accuracies are obtained using an ARMA modelling matrix.

AR order	MA order					
	3	4	5	6	7	8
2	95.8	98.0	96.1	95.5	96.6	96.9
3		85.7	88.0	87.4	85.2	83.8
4			92.6	92.4	92.7	89.9
5				81.2	80.1	82.1
6					77.0	84.3
7						78.2

coefficients at each decomposition level yields small error variances ensuring that features present in the wavelet domain are not lost with this modelling process. The classification performance obtained for both the cellular samples as well as the powder samples using the Mahalanobis distance classifier is encouraging and provides further motivation to use more elaborate data-driven schemes in future work such as multi-layer perceptrons or support vector machine classifiers. It is worth noting, however, that classification work using an SVM classifier scheme without adopting a wavelet decomposition step, but using four dimensional feature set based on amplitude and phase data has had a 89% success rate, at correctly classifying cancer cells (Withayachumnankul *et al.* 2005).

The adoption of a THz imaging system in a histopathology lab for the identification of cancerous cells, in conjunction with our classification methodology that is capable of differentiating between small differences in the transmission properties of the samples under study, is expected to provide significant advantages as current techniques are laborious, expensive and the results obtained are subjective to human interpretation. A further advantage of the adopted methodology is that it can be implemented in real-time so that measurements obtained with terahertz pulse imaging systems by practitioners from other disciplines can be displayed directly as feature maps in a manner that is transparent to the user. The work described will underpin the further proliferation of THz transient spectrometers in the biomedical, security, and pharmaceutical technology sectors.

9.5 Support vector machine applications in terahertz pulsed signal feature sets

Support vector machine (SVM) learning algorithms are sufficiently powerful to detect patterns within noisy signals. This Section introduces a frequency orientation component method to extract T-ray feature sets for the realisation of two- and multi-class classification using SVMs. Effective discriminations of ribonucleic acid (RNA) samples and various powdered substances are demonstrated as case studies. The development of this method is of important for in T-ray chemical sensing and image processing, which results in enhanced detectability useful for many applications, such as quality control, security detection, and clinical diagnosis.

As mentioned above, T-rays have promising potential both in *in vivo* and *in vitro* biosensing applications (Woodward *et al.* 2002, Mittleman *et al.* 1999, Mittleman *et al.* 1996) owing to (i) their non-invasive property, and (ii) the fact that biomolecules have rich resonances in the T-ray region (Withayachumnankul *et al.* 2005, Ferguson *et al.* 2002c, Siegel 2004). It is important to devise effective feature extraction methods to fully represent the different characteristics of these signals (Trier and Jain 1996). Signal processing methods are proposed for the current experiment. In the two experiments presented in this Section, input measurements are decomposed based on specific properties of their Fourier spectra (Wang *et al.* 2004a, Wang *et al.* 2004b). From the spectrum, a pair of specific values are extracted as features, which take the place of a large amount of spectral data. In doing so, the number of features is kept smaller than the number of observations to prevent overfitting. In this case, there is the added benefit of reduced computational complexity with low feature dimensionality.

The current experiments illustrate the potential of support vector machines (SVM) in pulsed signal recognition. The classification of two different types of RNA samples is of importance because it is a potential precursor for DNA sample classification in an *in vivo* environment (Fischer *et al.* 2005b). We also classify six different types of powder materials, to illustrate the validity of our approach to other applications, such as security. We describe the methods for T-ray pulse classification. The system's inputs are the measured raw T-ray signals, with the only preprocessing being a deconvolution in the Fourier domain. Feature extraction and classification are performed with the aim of achieving high classification accuracy. In this work, we found that an SVM kernel-based method can be directly applied to the specific features. The main advantage

of the current approach is that it is not necessary to design an elaborate feature extraction scheme (Kim *et al.* 2002), since even elementary feature extraction is sufficient for the SVMs to perform well. A pairwise classification scheme (Schölkopf *et al.* 1998) is applied to multiclass recognition problems, which leverages the good performance of binary SVMs to solve complicated multiclass problems. Gaussian and polynomial kernels have been found to give good performance for two-class and multiclass classification problems, respectively. Visual classification performance is also shown for the two dimensional features of binary and multiple classification, which gives useful demonstration of the effectiveness of the pattern learning problem and makes the performance of SVMs more understandable.

9.5.1 Terahertz data representation

This Section describes the measurement of T-ray pulsed responses. Terahertz pulsed imaging (TPI), or terahertz time domain spectroscopic functional imaging, is used for the current experiment to detect RNA samples. The complete T-ray imaging system results in T-ray waveforms for each position of the object in the x-y plane, therefore, a T-ray image can be built up pixel by pixel. The image formed in this fashion affords the T-ray pulsed response full information in both amplitude and phase (time delay). There are two target data sets: one is for RNA samples and the other is for powdered substances. The former contains two separate classes of biological specimen and the latter contains six various types of powdered materials. The representation of data measurement is detailed in the following subsections.

RNA data representation

Recently, it has been suggested that biological material can be detected by T-ray circular dichroism (TCD) spectroscopy, because many biomolecular crystals exhibit strong and specific absorption features in their dielectric spectra (Fischer *et al.* 2005b). Currently, the identification of the binding state of DNA is an especially interesting topic, which can be realised through applying T-ray techniques, i.e. loading a planar T-ray resonator with the sample material, in spite of the lack of characteristic absorption features in the T-ray region. In this experiment, it is investigated that frequency domain data from two different RNA polymer strands, polyadenylic acid (poly-A) and polycytidylic acid (poly-C) can be used for the recognition task with the potential to classify DNA.

Commercially available poly-A and poly-C potassium salts were used for the experiment (Sigma-Aldrich, product numbers P9403 and P4903), the data of which was measured at the University of Freiburg, Germany, and Fig. 9.15 illustrates the chemical structures of the polymers.

Figure 9.15. Chemical structure of RNA samples. (a) polyadenylic acid (poly-A) and (b) polycytidylic acid (poly-C). Potassium ions are applied to neutralize the negative charge of the backbone. After Fischer *et al.* (2005b)

The sample preparation has been described in detail by Fischer *et al.* (2005b), and the following steps are a short summary: (i) the carrier substrate was a commercially available cyclic olefin homopolymer slide (Greiner Bio-One) with transparent and dispersion-free properties in the T-ray frequency range; (ii) in order to achieve T-ray imaging, small liquid volumes were spotted by hand in a 4×4 array of alternating poly- A and poly-C on the slide. Each spot with a diameter of approximately 1 mm consisted of 2 μl of deionized water containing 0.2 mg material. The spots were then dried at room temperature and checked for homogeneity.

As a result, the terahertz image is illustrated in Fig. 9.16, which was achieved by terahertz time-domain spectroscopy imaging system based on free-space propagation and aperture-less focusing of the T-ray beam. Each pixel in the image represents the normalized peak values corresponding to Poly-A and Poly-C. The sample consists of a 4×4 array of spots. Two of the spots were removed from the substrate in order to identify the orientation of the substrate in the image. The spot of Poly-A is shown at the top

left corner of the image, with weak transmission, compared to the spots of poly-C. The positions of poly-A and poly-C subimages are labelled in the diagram to the right of the picture.

Based on the positions of poly-A and poly-C at the terahertz image, we select 8-neighbor pixels around center position from each spot for the signal post-processing and classification. The pixels lying on the boundaries of each class are excluded from training and test vectors due to the difficulties involved in manually assigning the desired values for pattern recognition.

Powder data representation

Another topical application of T-rays has been the classification of powdered samples as a basis technology in substance detection for security (Ferguson *et al.* 2003). A general question is to explore the ability of T-ray spectroscopy to detect different densities, thicknesses, and concentrations of specific powders. To investigate this problem, we conduct a preliminary exploration of different powder recognition tasks with 2 mm thickness for six different powdered substances and their holder. They are: sand, talcum, salt, powdered sugar, wheat flour, and baking soda.

The sample holder is shown in Fig. 9.17; this holder has an ability to accurately control the thickness of the powders. There are two Teflon blocks, which can be separated under control. They are mounted on a manual translation stage to provide the required gap of 2 mm, where a plastic bag containing the powdered substances can be inserted between the Teflon blocks. This procedure guarantees a relatively consistent powder density and accurate control over the powder thickness of 2 mm.

A traditional T-ray imaging system is used to detect T-ray responses based on the THz-TDS technique. There is an x-y translation stage involved for mounting the Teflon sample holder and fixing its position in the T-ray beam. At T-ray frequencies, Teflon is dispersionless. As Teflon has a very low absorption coefficient, there is minimal distortion while the T-ray pulse propagates through the holder. A 2D T-ray image of the sample can be obtained after inserting a powder sample. This image allows the effects of different scattering paths and minor variations in powder thickness and density to be observed. In general a 1D image is sufficient for substance detection purposes, and 50 pixel responses (with a pixel spacing of 100 μm) can be acquired in under 30 minutes.

NOTE:
This figure is included on page 186 of the print copy of the thesis held in the University of Adelaide Library.

Figure 9.16. T-ray transmission image of poly-A and poly-C RNA samples on a biochip substrate. It is illustrated that there exists stronger absorption in poly-C compared to poly-A. Each spot contained 200 μg of either poly-A or poly-C in alternating order, as indicated in the diagram to the right. The colour scale indicates the normalised peak values of the two RNA samples. After Fischer *et al.* (2005b)

9.5.2 Terahertz feature extraction

This Section describes the classification system designed to assess the potential of SVMs in T-ray pulsed classification. There are two target data sets, which need to be separated by SVMs: one is for RNA samples and the other is for powder samples. The former is to classify two classes of objects and the latter is to separate six various types of powdered materials. RBF kernels and polynomial kernels are applied for statistical feature mapping. Signal processing is applied to track the key features of training vectors for different classes of signals.

Feature extraction via frequency orientation components

The way to extract specific feature vectors in the frequency domain is realised by taking the Fourier transform after deconvolving measured signals with a reference pulse (Ferguson and Abbott 2001a). The Fourier transform produces complex-valued spectra, containing both phase and magnitude information. The magnitude and phase at certain key frequency components constitute pairs of feature subsets on which the classification is based. An important advantage of this approach is the small dimensionality of feature vectors, allows the features to be directly extracted from pulsed

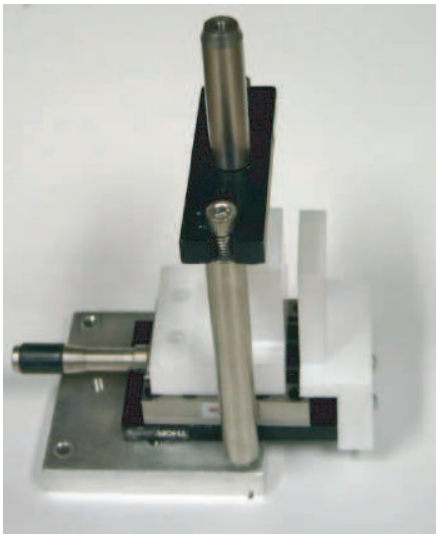


Figure 9.17. Photo of a teflon sample holder for measurement fixed thickness of powdered samples. One of the teflon blocks is fixed, while the position of the other is controlled using the manual translation stage. The gap between the two blocks may be adjusted to allow 2 mm thicknesses of powder to be considered.

responses with relatively low computational complexity. This obviously reduces the computational resource requirements, which is attractive for possible hardware-based implementations. Further, the sparse features help us avoid the over fitting problem (Guyon *et al.* 2002, Withayachumnankul *et al.* 2005). The input vectors occupy a size $M \times L$ matrix, where M is equal to the number of input vectors (training vectors) and L is the dimensionality of each feature vector. If the limited measured dimensions of the training subsets M is smaller than large scale of time series features L, it will cause difficulty in correctly assigning labels to target samples. Fast and sparse features overcome the computational disadvantages of SVMs.

Kernel selection and parameter tuning

Following feature extraction, the kernel operation is performed on the calculated features in an SVM. An implicit nonlinear transformation Θ is used to map input pattern $\phi(\mathbf{x})$ into a higher-dimension, yet linear, space. In the case of a Gaussian kernel, this transformation is related to a Gaussian function $K(\mathbf{x},\mathbf{y}) = \exp(-\gamma \parallel \mathbf{x}-\mathbf{y} \parallel^2)$, where γ is the Gaussian kernel width parameter. Accordingly, an SVM classifier needs to be applied to produce learning vector patterns in two dimension feature space. The scatter plot for learning vectors using a Gaussian kernel is illustrated in Fig. 9.19, which considers a two-class pulsed signal classification problem for the recognition of two types

of RNA samples—poly-A and poly-C. It serves as a useful comparison with plotting machine learning realised via a polynomial kernel for classification tasks on multiclass powder classification. The feature vectors consist of magnitude and phase, which are plotted in a two dimensional plane: x axis is labelled by magnitude and y axis labelled by phase.

The parameters that need to be tuned in the SVMs include the kernel parameters and penalty parameter C (Weston *et al.* 2003, Hastie *et al.* 2003). The use of a *validation data set* can be viewed as a straightforward method for tuning. This method reserves from the training data set, a number of features for validation (Kim *et al.* 2002), which are used to gauge the performance of the classifier trained on the remaining training data. In our experiments on the RNA data, we have chosen the validation data set approach for the RNA data. However, the use of validation sets becomes infeasible when the number of feature vectors in the data sets is small. This is the case for the powder classification experiments. In that case, we have chosen to forego the validation set, and instead use the test error as a direct way of comparing against different parameter values. In spirit, this approach is akin to an exhaustive search for optimal SVM parameters, while it is reasonable when there is a paucity of data and few parameters to determine.

9.5.3 Performance assessment of classification

Cross-validation methods (Bengio and Grandvalet 2004, Cai and Li 2005) and a leave-one-out (LOO) (Fukunaga and Kessell 1973, Fukunaga and Hummels 1989) estimator within the deconvolved T-ray data set are utilized to provide a nearly unbiased estimate of the prediction error rate. The performance of classifying the RNA samples is evaluated using 8-fold cross-validation, while the powdered material classification is validated using LOO. The data set of RNA is divided into eight subsets of approximately equal size. Sequentially, each subset is tested using the classifier trained on the remaining subsets. The results from the eight runs are averaged to provide a statistical estimate of the classifier performances. To tune the parameter C, we use small-and-separate validation sets drawn from the test subsets, with the remainder of the test subsets used for testing the classification performance.

In the approach outlined above, each RNA pixel label is predicted once so the cross-validation accuracy is the percentage of data which are correctly classified. Similarly,

LOO evaluates each unknown feature vector and then produces a basis to evaluate classifier designs for powder classification (Fukunaga and Hummels 1989, Yin *et al.* 2007f). Therefore, LOO accuracy is also the percentage of correctly classified data sets. The reason why LOO is used instead of 8-fold cross-validation for the powder experiment is due to the relatively small number of measurements for the different powders. With such a restriction, LOO is preferred as the overall classification experiment is averaged over more runs. In this experiment, accuracy of classification is used as the quantity for assessing the performance of all the classification tasks, and it is calculated by:

$$\text{accuracy} = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FN} + N_{TN} + N_{FP}} \tag{9.1}$$

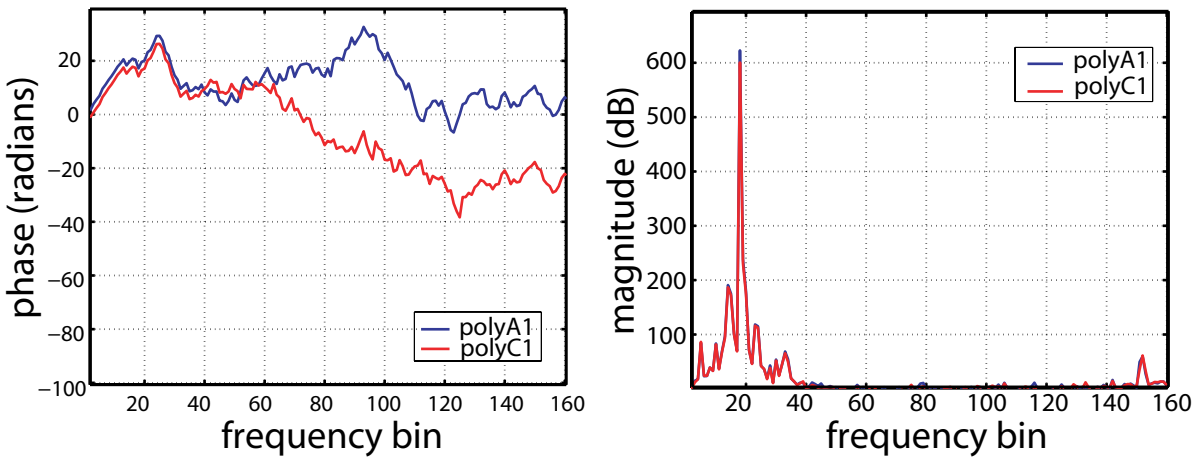
where a true positive is labelled by N_{TP}, a true negative is labelled by N_{TN}, a false positive is labelled by N_{FP}, and a false negative is labelled by N_{FN}.

9.5.4 The Fourier spectrum analysis

As mentioned above, the detection procedure works on a modified transmitted time-series of T-ray pulsed responses. For isolation of system response, a pulse measured from an empty substrate and an empty holder is employed as reference and is deconvolved from the measured signals of RNA samples and powdered materials, respectively, for the removal of the system response (Ferguson and Abbott 2001a, Duvillaret *et al.* 1996). The specific features relevant to magnitude and phase are extracted from the RNA and powdered substances data, via a Fourier transform. The details are described in the following subsections.

The Fourier spectrum analysis for the classification of poly-A and ploy-C T-ray pulses

RNA data occupies a size equal to $60 \times 50 = 3000$ pixels. For each pixel, the number of time samples is 350, which was accordingly truncated at 175th frequency bin corresponding to frequency of 4 THz because of the symmetry of spectrum. The 3000 pixel data set consists of background data information—a TOPAS substrate image, and target object data sets—poly-A and poly-C image data. The population of pixels belonging to the poly-A and poly-C classes is 48 for both classes. In order to obtain reduced dimensions of feature subsets and make them discriminable for the different classes, the magnitude and phase values of the pulse responses are first calculated, and then those values corresponding to the frequency with the greatest magnitude (i.e. strongest



(a) (b)

Figure 9.18. Illustration of T-ray spectra of RNA. (a) Displays two obvious separate phase curves for two classes of spectra—poly-A and poly-C, which are obtained by down sampling the Fast Fourier Transform (FFT) of T-ray pulses with linear extrapolation. (b) The plot of magnitude versus different frequency bins with largest magnitude at frequency bin 19.

response) are used as the features. This process extracts a 2-dimensional feature vector from the full spectral data with 350 non-redundant dimensions.

Fig. 9.18(a) displays two obvious separate phase curves for two classes of spectra—poly-A and poly-C, which are obtained by Fast Fourier Transform (FFT) of the T-ray pulses with linear extrapolation. Fig. 9.18(b) is the plot of magnitude versus different frequency steps with cutoff frequency at the 175^{th} frequency bin. It is observed that the magnitude at frequency bin 19 reaches maximum value. Hence, we select frequency bin 19 as the key frequency of interest and use the corresponding phase and magnitude pair as the extracted features, which form the input to the SVM. A Gaussian kernel is used for the final feature mapping from a non-linear feature space to a linear one. Accordingly, we apply the SVM training algorithm to produce a learning vector pattern in the two dimensional feature space (magnitude and phase form the axes), which is illustrated in Fig. 9.19. SVMs with the width parameter of Gaussian kernel λ of 0.003 are trained by feeding 6 to 42 training vectors selected randomly from 48 patterns from each class. The orientation frequency is selected at the 19th frequency bin in all cases. The background colour shows the shape of the decision surface. Dark blue regions represent the class belonging to the poly-C sample labelled by -1; and light blue regions indicate the class related to poly-A sample labelled by 1. Separating hyperplanes for

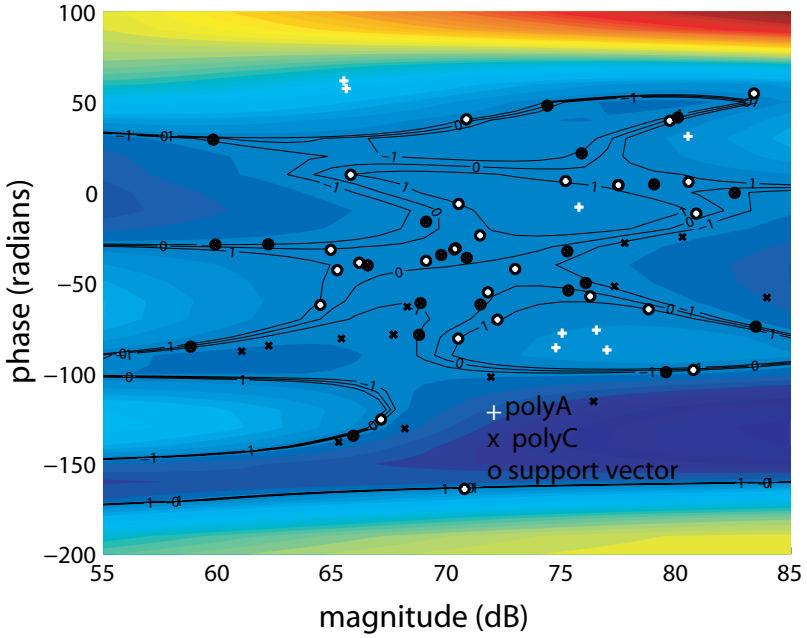


Figure 9.19. An illustration of binary classification for the recognition of RNA samples. . A Gaussian kernel is used for mapping the training vectors to a 2D feature space. The penalty parameter C is set to infinity and the width parameter of Gaussian kernel λ equals 0.0006. The background colour shows the contour shape of the decision surface. The learning vectors are approximately separated via applying Gaussian kernel for mapping. This resultant mapping shows good discrimination of two-classes RNA samples.

two classes are indicated by 0. The circles represent the calculated support vectors. Compared to the training vectors, the number of support vectors are reduced, which takes on an important role in achieving the ideal shape of hyperplanes and facilitating computation of the classification algorithm. In this case, machine learning for two-class samples—poly-A and poly-C denoted by white '+' and black '$\times$' are approximately separated by their own boundary lines though there is a little overlapping. Detailed experimental results about classification accuracy are analysed in the next Section, after 200 random selections of training vectors are fed to the SVMs.

Fourier spectrum analysis for multiclass classification

The image statistics of powders consist of $6 \times 50 = 300$ pixels. For each pixel, the number of time samples is 400. Fig. 9.20 shows the phase and magnitude plots in the frequency domain from one pixel of salt image data, with a cutoff frequency equal to 4 THz. It is obvious that the subimage at the bottom of Fig. 9.20 shows a sharp change

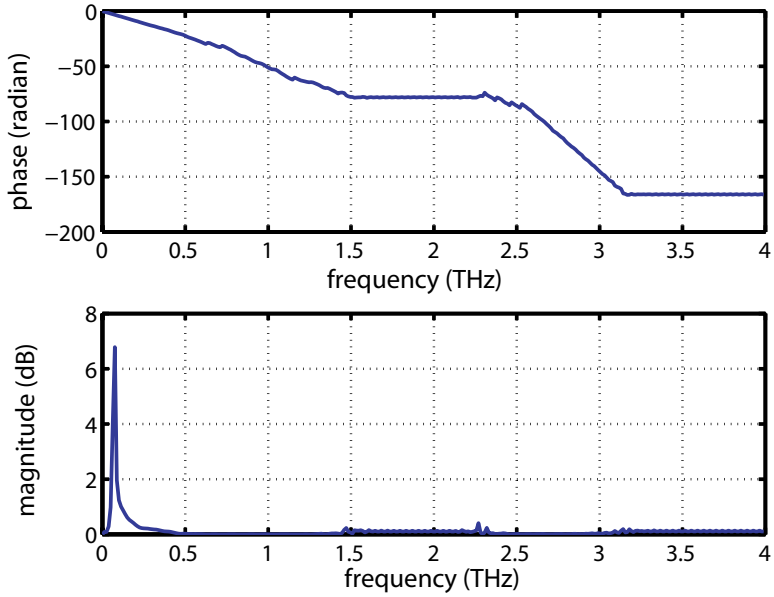


Figure 9.20. Illustration of Fourier spectrum. The top subimage shows phase versus terahertz frequency and the bottom subimage shows magnitude versus terahertz frequency.

of magnitude at the second frequency bin. Accordingly, with the number of training subsets of 49, we plot the phase and magnitude based on two classes learning vectors and multi-classed learning vectors, by applying the pair of Fourier features at the second frequency bin, then we produced the learning vector pattern for two-classes and multiclass recognition, shown in Fig. 9.21 and Fig. 9.22, respectively. For powder classification, a polynomial kernel is employed for optimal classification performance. The two-class powder samples, sand and salt samples, are linearly separable, with circles describing the calculated support vectors, which decide the linear optimal hyperplane between two classes. The solid lines above and below the hyperplane depict the ± 1 range along the separating surface. The small number of support vectors greatly reduces the computational burden of the classification task. In two-class classification, the penalty parameter C was chosen to be 10 and the polynomial kernel degree equals 3. All these parameters are tuned experimentally.

For the multiclass case, a polynomial kernel with degree of 3 is applied for linear mapping, with a truncated terahertz frequency of 4 THz. The penalty parameter C is set to 100. The relevant decision functions for the pairwise approach are shown in Fig. 9.22 with the number for the recognition of the various decision surfaces corresponding to the different pairwise classes. The summing up of the pairwise votes yields the borders easily. The small red region at the left hand outlined by border line 5 is an undecided

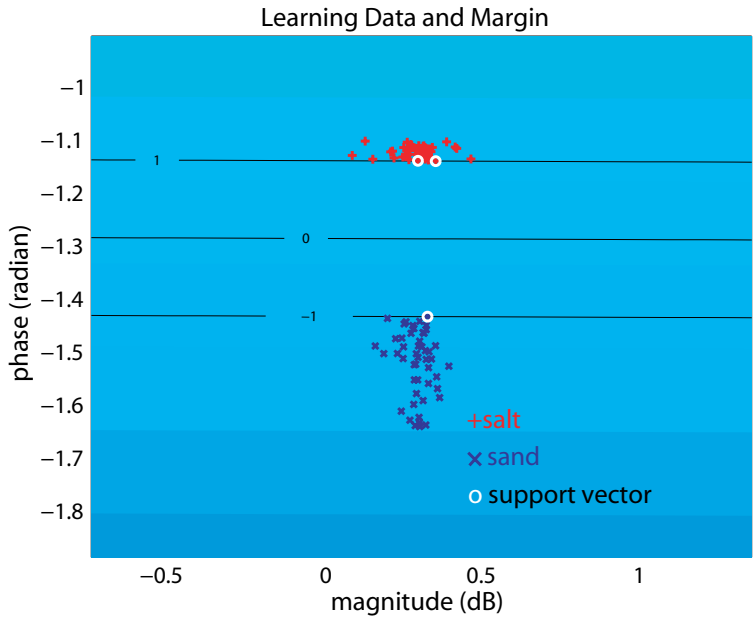


Figure 9.21. Plot of learning vector pattern. The classifier is trained with 49 pixels from two-class powder samples, sand and salt samples, with phase and amplitude as a pair of coordinates. Polynomial kernel with degree of 3 is applied for linear mapping. The penalty parameter C is set to 10.

class region. This is because the maximum number of votes (top scores) in the region is smaller than $(k-1) = 5$. The six classes are well clustered and therefore easily separated. The undecided class region has no relevance for the class decision. It is obvious that the number of the support vectors for the single decision is small, which results in a fast adaption and better boundary shape to partly compensate the computation increase brought on by needing to repeat $(k-1)/2$ calculations for multiclass classification. The support vectors as the 'most important' data points are identified with extra circles.

9.5.5 Resultant classification performance

To verify the effectiveness of the proposed method, resultant classification were performed on the extracted features of T-ray pulsed data related to RNA samples and several types of powder substances. 'LIBSVM' (Chang and Lin 2001) and with 'SVM and Kernel Methods Matlab Toolbox' (Canu *et al.* 2005) are chosen as the toolboxes for our experiments with two- and multiclass classification, respectively.

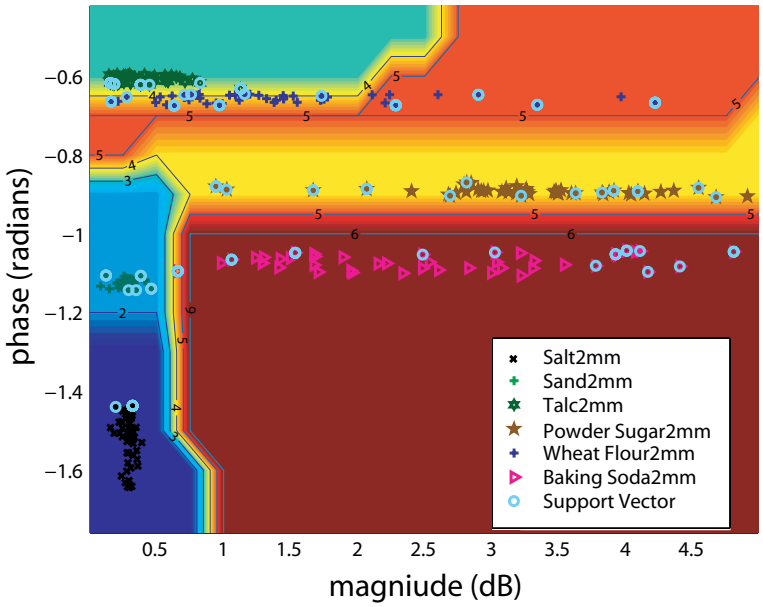


Figure 9.22. Learning vectors for the six-class examples are plotted. It is for the illustration of the linear decision function between the pairs of classes by applying a polynomial kernel for mapping with a degree of 1. There are 49 pixels selected randomly from each of six classes of powder samples. The small red region on the left hand side remains undecided.

RNA classification via Fourier coefficient spectrum

For classification of RNA samples, Gaussian kernels are applied because of the preferred classification performance that fit the RNA data well. All the classification runs are performed using a 2.4 GHz Pentium4 CPU. The average time spent classifying the two-classes RNA samples is 2.74 seconds for 18 data sets.

To evaluate the effect of the Gaussian kernel on the RNA sample classification, suitable values of C and λ are evaluated via parametric search using separate validation sets. After training, the final error rate, the number of support vectors and the elapsed time are compared. In the training phase, the training vectors are randomly selected from a given proportion, varying from 1/8 to 7/8, of the input population of 48 pixel responses from each RNA class. The SVM parameter (C) is tuned by the remaining 1/16 to 7/16 of input data as validation vectors and tested with the last small-and-separate subsets of 1/16 to 7/16 of available pixel responses. In principle, a similar procedure can be applied to the tuning of parameter λ—in this Thesis, we illustrate the tuning of C for brevity, since we have found the classification performance to be less sensitive on the choice of λ. As discussed in Section 9.5.4, the key frequency is selected at the 19th

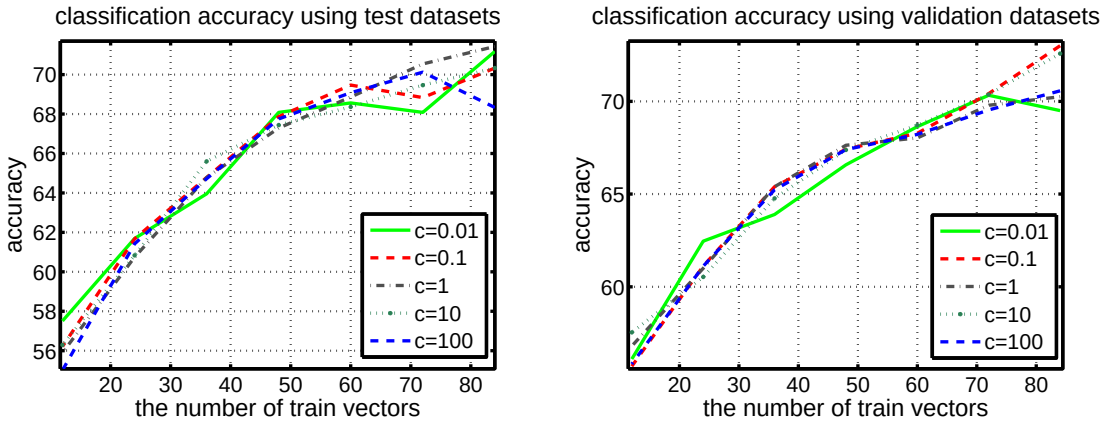


Figure 9.23. Plots of classification accuracy regarding RNA samples. (a) Plot of classification accuracy versus the number of input test vectors, corresponding to different values of parameter C. (b) Illustration of the validation of classification accuracy, via the plot of classification performance versus the number of input validation vectors, corresponding to different values of parameter C.

frequency bin in all cases. In order to achieve effective classification, the repeated selection of test and training vectors are conducted, and the number of repeated selection approaches 200. The highest classification performance was obtained for the penalty parameter $C = 1$ and $\lambda = 0.003$, with a classification accuracy of 72%.

It was found that the classification accuracy is similar throughout the range of values for C, from 0.001 to 10^4, in steps of 1 on a log scale. The classification accuracy is improved with an increased number of training vectors, which is to be expected. Fig. 9.23(a) and (b) show classification performance using our algorithm versus different sizes of test and validation data sets, respectively. The various value of C are all plotted for direct comparison. It is clear that the two subfigures show the similar behaviour, which implies that the validation and test sets do not exhibit very different classification characteristics. In Fig. 9.23(a), the curve, related to C of 1, gives best performance, especially when the number of training vectors is in the range from 48 to 84, though in Fig. 9.23(b), the corresponding curve shows a slightly weaker classification accuracy compared to the others.

The number of the computed support vectors is roughly one-third fewer than the number of training vectors. A small number of SVs is desirable for implementation since it directly determines the computational complexity of the automatic classification task.

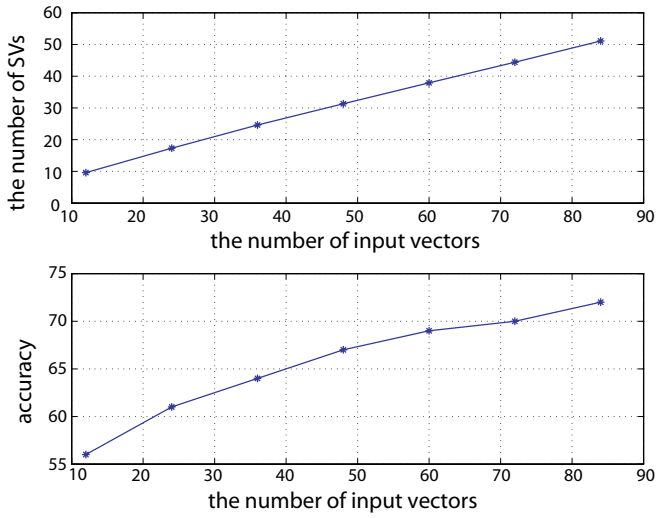


Figure 9.24. Illustration regarding the variation with respect to the number of SVs and classification accuracy. There are plots against the increment of the number of the input training feature sets in the classification of RNA data, with $C = 1$ and $\lambda = 0.003$.

Fig. 9.24 below shows the variation of the number of SVs with the increment of the number of the input training feature sets in classification of RNA data. The number of the input training RNA samples varies from 12 to 84, with a step size of 12. The corresponding number of SVs, shown in Fig. 9.24(a) shows an almost linear increment from 9.6 to 51.4. Similarity, the classification accuracy, shown in Fig. 9.24(b), varies with a range from 56% to 72%. It can be approximately viewed as a linear increase with a 2% improvement in accuracy for every 10 additional training vectors.

The feature extraction method realised by selecting key frequency components is particularly attractive when the input vectors come from large data sets. However, it should be noted that the current classification results for RNA samples is limited by laser fluctuation occurring in the measurement procedure between an RNA sample and the substrate reference. In addition, the RNA data is measured on a very thin layer of the substrate (around 40 μm). The propagation delay is often less than 1 sampling period, which makes model fitting difficult.

Multi-class powder classification via Fourier coefficient spectrum

Table 9.9 to Table 9.11 show the multiclass classification performance via applying a pairwise classification method, with the application of polynomial kernels, p from 1 to 3, and varying penalty parameter C from 0.1 to 1000 with step of 1 in a log scale, on Fourier spectral features. Elapsed time of the SVM testing and the number of SVs are

also shown in the tables. A leave-one-out estimator is used for the training and testing of the SVMs. The total size of analyzed vectors of each target class is $50 \times 400 = 20000$ before feature extraction. For training, $49 \times 6 = 294$ pixels from all the classes are input to SVMs. The remaining 1 pixel from each class is used to test the SVMs. The SVM experiments are repeated 50 runs. Therefore, the test lapsed time indicates the 300 runs of SVMs testing time. All the powder classification runs are performed using a 1.66 GHz Intel dual-core CPU.

Table 9.9. Varying penalty parameter C and the polynomial kernel p=1 are implemented for the parameter calculation regarding SVMs. Classification rates (%), number of SVs and elapsed time are illustrated. The Fourier spectral features were extracted. Here, 97.96% and 57.14% classification accuracy correspond to the number of the correct classified test vectors equal to 48 and 28 with 1 and 21 test errors, respectively.

	each class name and classification accuracy (%)						♯ of SVs	elapsed time (s)
C	salt	sand	talcum	sugar	flour	soda		
0.1	95.92	83.67	97.96	83.67	57.14	34.69	566	35.6
1	100	100	93.88	95.92	67.35	89.80	334	23.9
10	100	100	95.92	100	83.67	93.96	112	15.0
100	100	100	97.96	100	100	97.96	64	12.6
1000	100	100	97.96	100	100	97.96	46	11.9

Table 9.10. Varying penalty parameter C and the polynomial kernel p=2 are implemented for the parameter calculation regarding SVMs. These parameters involving classification rates (%), number of SVs and elapsed time are illustrated for powder classification. Here, 97.96% and 67.35% classification accuracy correspond to the number of the correct classified test vectors equal to 48 and 33 with 1 and 16 test errors, respectively.

	each class name and classification accuracy (%)						♯ of SVs	elapsed time (s)
C	salt	sand	talcum	sugar	flour	soda		
0.1	100	100	95.92	95.92	67.35	97.96	374	20.0
1	100	100	93.88	100	77.55	97.96	120	17.0
10	100	100	97.96	100	93.88	97.96	78	15.0
100	100	100	97.96	100	100	97.96	59	13.9
1000	100	100	97.96	100	100	97.96	50	12.8

The maximal vote is selected as a winner for the recognition of a target. The averaged classification accuracy in relation to each powder sample is shown in the three tables based on the various penalty parameter C and the degree p of three polynomial kernels. Table 9.9 to Table 9.11 shows that the classification accuracy versus the penalty parameter C of 0.1, 1, 10, 100 and 1000 for their corresponding kernel parameter value. It is obvious that, with the increase of C and polynomial kernel of p, the classification

Table 9.11. Varying penalty parameter C and the polynomial kernel p=3 are implemented for the parameter calculation regarding SVMs. They are classification rates (%), number of SVs and elapsed time for powder classification. Here, 97.96% and 71.43% classification accuracy correspond to the number of the correct classified test vectors equal to 48 and 35 with 1 and 14 test errors, respectively.

	each class name and classification accuracy (%)						♯ of SVs	elapsed time (s)
C	salt	sand	talcum	sugar	flour	soda		
0.1	100	100	95.92	95.92	71.43	97.96	157	19.2
1	100	100	95.92	100	81.63	97.96	89	16.7
10	100	100	97.96	100	100	97.96	71	15.9
100	100	100	97.96	100	100	97.96	54	14.7
1000	100	100	97.96	100	100	97.96	51	14.9

accuracy improves, while the averaged number of support vectors and the elapsed time both reduce. When the value of p is set to 1, 2, and 3, the classification performance is satisfactory when C is set to be greater than 100, 100, and 10, respectively, with the accuracy being a perfect 100% for all but two powders: talcum and soda samples, which scored 98%. The p of polynomial kernel of 3 is suitable for classification of terahertz powder data, at which the C can be set to a small value—this is important because if C is too large, there is the risk of over-penalising training error, and hence over-fitting the SVM. It should be noted that in the three tables, the 97.96% and 57.14% classification accuracy correspond to the number of the correct classified test vectors equal to 48 and 28 with 1 and 21 test errors, respectively. In this case, 1 test error reduces the classification accuracy by 2.04%. Averaged classification accuracy over the six classes is 99.32%.

It should be noted that the relatively few SVs when setting $C >= 10$, with an average number of 65, are required when compared to the large number of input training vectors equal to $50 \times 6 = 300$; once again, having fewer SVs reduces the computational load of the testing phase. The average elapsed time is measured for each class classification, which is approximately equal to 9 seconds when using a polynomial kernel with degree from 1 to 3.

For comparison with the proposed SVM-based powder classification performance, the Mahalanobis classifier and a k-means classifier are selected to achieve supervised and unsupervised classification, respectively. For the Mahalanobis classifier, an iterative algorithm is employed to select two key frequencies among the first 50 frequencies (Yin *et al.* 2007e). Half of pixels (25 pixels) from six types of powders at 2 mm are used

to train the classifier, and the remaining part (25 pixels) at 2 mm are used to test it. The classifier testing time is 682 s after 200 run testing on the same hardware. For the unsupervised *K*-means classifier, the exact same feature set consisting of amplitude and phase at the second frequency component is used as the input, and the number of clusters is set to be 6. The algorithm is initialised using randomly selected cluster centres; the training is repeated 100 times and the lowest mean square Euclidean distance clustering is used as the final result. Under these conditions, the unsupervised classification accuracy is 75%.

9.5.6 Conclusion

This Section describes an SVM-based T-ray pulsed signal classification method to estimate the potential of SVMs in the classification of RNA samples, poly-A and poly-C, and the classification of six types of powder materials. Pairwise classification is employed for the multiclass powder samples. The suggested method is to put the original T-ray pulses into SVMs, which does not involve any external feature extraction scheme except for the adoption of the normalization and fast Fourier transform for signal and spectrum analysis. The principle for this ability is that SVMs allow processing of sparse features for machine learning in low-dimensional feature spaces. The validity of using Gaussian and polynomial kernels is supported by effective classification performance of the above two feature extraction procedures.

It is observed that SVM implementation is not too onerous on hardware, as long as training is done off line. The testing is very fast and quite low on computational resource requirements. Since the classification performance is strong, implementing a SVM-based classification system is both feasible and worthy of consideration.

Future studies for a SVM-based classification system will investigate further signal processing techniques and statistical modelling (Schürmann 1996, Schalkoff 1992) based methods for different feature extraction of T-ray pulses. Autoregressive modelling (Therrien and Oppenheim 1992) can be employed as a technique to achieve the decomposition of large number of measured time samples. Absorption coefficients and refractive index are also suggested to be used as good choices for the key features extraction, especially while analyzing the RNA spectrum data affected by obvious etalon artifacts. Meanwhile, more experiments related to DNA need to be carried out for the exploration of different feature configuration and different multiclass classification

methods. Additionally, the powder classification experiment was conducted based on the THz measurements with same thickness of powder samples to train and test a classifier. It is in hope that different sample thickness can be taken advantage of for training and testing purposes.

9.6 Chapter summary

This Chapter uses specific THz experiments to validate feature extraction and classification algorithms described in previous Chapters, containing the central results of THz pattern recognition.

These THz experiments take advantage of T-ray time domain techniques, which allow the measurement of T-ray pulsed responses in both amplitude and phase (time delay). The T-ray time domain response is obtained via terahertz pulsed imaging (TPI).

Section 9.2 highlights the ability of the wavelet SURE soft threshold to shrink the THz signals dominated by system noise. It is validated via THz classification experiment of different powder specimens. In this experiment, FFTs are used as a feature extraction tool and the Mahalanobis distance classifier is employed for classification. Three optimal frequency components are selected via an iterative algorithm and the resultant classification performance is satisfactory.

Section 9.3 explores ability of three different system identification algorithms in the discrimination of three powder state materials: lactose, mandelic acid, and dl-mandelic acid. Three signal processing algorithms regarding system identification techniques are adopted to improve on the signal to noise ratio of the calculated spectra, and to evaluate the complex insertion loss function of the samples under study. Complex insertion loss functions of the samples as important experiment parameters are evaluated for each processing technique. Compared to a subspace algorithm, a Mertz apodization window for ratioing the sample performs well in terms of smoothness of the output CIL, and shows correct identification of the absorption bands. However, after dispersion, discrimination matrices are applied for the evaluation of different sample discrimination, the standard ratioing procedure are inferior with respect to noise than those obtained by the other two identification algorithms. Wavelet-packet technique has been shown to be particularly suited to the identification of systems with several spectral resonance features.

Section 9.4 introduces a parametric modelling procedure. The AR and ARMA modelling of wavelet subband coefficients are applied for an effective feature extraction algorithm. We discuss the success of the proposed algorithm in classifying human bone osteoblasts (HBO) against human osteosarcoma cells (HOS) as well as at differentiating between six types of powder samples. Our positive results motivate future exploration of other classes of cancer cells via THz radiation. THz pulse transients of different powder samples illustrate the ability of THz to classify different substances, which potentially have important applications in the pharmaceutical and security industries.

Section 9.5 illustrates support vector machine learning algorithms that are sufficiently powerful to detect patterns hidden inside noisy biomedical measurements. A frequency orientation component method is applied to extract T-ray feature sets for the application of two- and multi-class classification using SVMs. Effective discriminations of ribonucleic acid (RNA) samples and various powdered substances are achieved. The development of this method has become important in T-ray chemical sensing and image processing, which are useful to enhance detectability for many applications, such as quality control, security detection and clinical diagnosis.

With every technological advance that has opened up new areas of the electromagnetic spectrum, there has been born a wealth of industries to apply that technology for the advancement of mankind—such is the promise of the THz regime (Abbott 2000). This Thesis presents significant and novel research on two parallel fronts towards pattern classification and computed tomographic reconstruction. In addition to the pattern classification (identification) system described in previous Chapters, from the next Chapter, this Thesis contributes to the field by developing processing algorithms to extend THz imaging capabilities to new application domains. Chapter 10 introduces the THz computed tomography, using a broadband THz source as an example, to illustrate the ability THz radiation to be used for computed tomographic reconstruction, in a similar way to X-rays.

Chapter 10

Terahertz Computed Tomography

COMPARED to simple planar imaging, computed tomography offers a more advanced mode of imaging, offering multiple viewing angles and internal structural information. It can be used to probe the nonlinear relationships between the signal strength, depth, and material optical properties. For terahertz computed tomography, a single point of a sample boundary is illuminated and the sample can be viewed at multi angles via a rotation stage connected to a linear stage on which the sample is mounted. Data is collected at the different positions and rotation angles around the boundary of the moving sample via a photodetector or a CCD camera. These THz measurements are then combined in a tomographic scheme, for reconstruction of intrinsic target contrast. Quantities used in reconstruction can be the absorption coefficients, refraction indices, or time domain parameters, such as time delays.

10.1 Brief review of THz imaging application

Terahertz Computed Tomography (THz-CT) is a form of optical coherent tomography, which offers a promising approach for achieving non-invasive inspection of solid materials, with potentially numerous applications in industrial manufacturing and biomedical engineering. While one- and two dimensional applications with time-domain terahertz spectroscopy have been well demonstrated in the past (Ferguson and Abbott 2001a, Galvão *et al.* 2003), the ability to non-destructively probe the inner three dimensional structure of optically opaque structures is less well studied. There has been a relative scarcity of terahertz tomography work in the literature. Currently, most THz pulsed imaging (TPI) schemes are based on the extraction of material parameters, which are calculated based on (i) time-of-flight measurements in reflection and transmission geometry or (ii) the variation in amplitude or phase of the transmitted waveform at certain frequencies (Mittleman *et al.* 1996). THz tomographic reconstruction is another interesting application of T-ray imaging. It exploits methods from geophysical, radar and optical diffraction techniques (Ferguson *et al.* 2002b). It is capable of mapping the 2D and 3D distribution of scattering objects, but with a constraint to image the shape profile of the target object. At present, three dimensional (3D) T-ray CT imaging, described in Ferguson *et al.* (2002b), has been developed based on coherent T-ray detection techniques, which allows the imaging of internal structure and extracts frequency dependent properties in three dimensions. Along with a dimension in time, THz-CT offers the possibility of rich 4D data sets describing the target sample.

10.2 Methodology of computed tomography

The advent of computed tomography (CT) in 1972 was a milestone in the history of diagnostic medicine (Kak and Slaney 1988). A tomographic slice, or the Radon transform of an object, is a cross-sectional integration of some property of an object, such as the attenuation coefficient in X-ray computed tomography. These measurements are collected via transmission or reflection of radiation sources set up to illuminate the object from many different angles. Computed tomography is based on the assumption that the image to be reconstructed belongs to an identifiable ensemble of similar images (Hanson and Wecksung 1983). A number of algorithms for inverting these measurements back into a two-dimensional image exist, but perhaps the most popular is the filtered back projection (FBP) method.

The filtered back projection algorithm, is often used in the calculation of the inverse Radon transform (Kak and Slaney 1988, Ferguson *et al.* 2002b, Yin *et al.* 2007e), which reconstructs a target object. Typically, the quantities obtained are the object's frequency dependent optical properties, such as the refractive index and absorption coefficient. Normally, a filtered back projection algorithm begins with a collection of *sinograms* obtained from projection measurements. A sinogram is a collection of points in a projection space, see Fig.10.1(b), which reflects projection geometric paths in an image space, see Fig. 10.1(a). A sinogram is generated using Radon transform data, denoted below by $\Re$ and is simply a collection of the projections at all the projection angles. A 1D projection operation at each projection angle is a linear integral of the image intensity along projection offset ξ, which satisfies the following equation:

$$s(\xi,\theta) = \int o(x,y)\mathrm{d}\xi = \Re(o) \tag{10.1}$$

where all points on projection offset ξ satisfy the equation: $x\cos\theta + z\sin\theta = \xi$ and $o(x,y)$ denotes the measured image intensity of a target object, which is a function of pixel position in an x and z plane.

The filtered back projection algorithm for terahertz CT reconstruction (Kak and Slaney 1988) is expressed as follows:

$$I(x,y) = \int_0^{\pi}\left[\int_{-\infty}^{\infty} S(\theta,\beta)|\beta|\exp[i2\pi\beta\xi]\mathrm{d}\beta\right]\mathrm{d}\theta \tag{10.2}$$

where $S(\theta,\beta)$ is the spatial Fourier transform of the parallel projection data, defined as

$$S(\theta,\beta) = \int_{-\infty}^{\infty} s(\theta,\xi)\exp[-i2\pi\beta\xi]\mathrm{d}\xi, \tag{10.3}$$

here, $s(\theta,\xi)$ is the measured projection data, β is the spatial frequency in the ξ direction. More details associated with the back projection algorithms are presented in Appendix B. It should be noted that the operation of the ramp filter $|\beta|$, as illustrated in Eq. (10.2), is equivalent to a differentiation followed by a Hilbert transform, which introduces a discontinuity in the derivative of the Fourier transform at zero frequency. It will cause an ill-posed inverse problem. This is the reason why a wavelet based algorithm will be introduced in Chapter 11.

In the next Section, we will introduce the THz CT imaging setup and illustrate the image mode for better understanding the THz CT reconstructions.

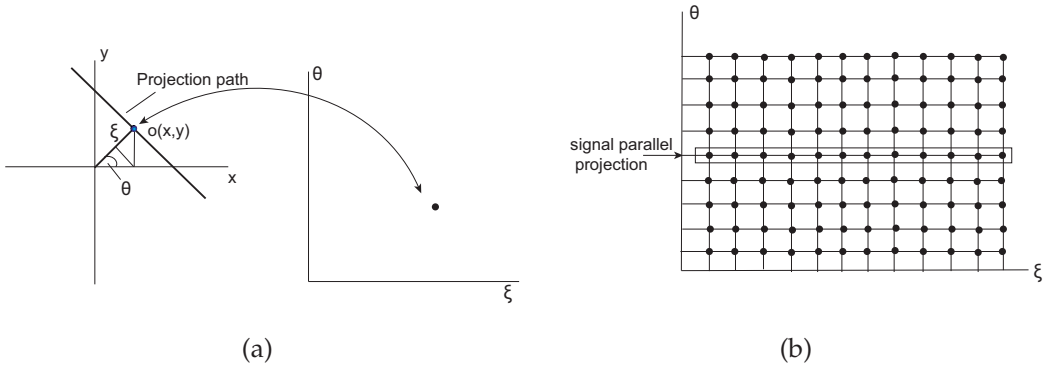


Figure 10.1. Illustration of general scheme for computed tomography. (a) A geometric projection path in an image space is a point in a related projection space; and (b) A projection space sample pattern.

10.3 Brief introduction to terahertz imaging for CT

The CT data that we will use in Chapters 11 and 12 is due to (Ferguson *et al.* 2002b) and the hardware used is briefly reviewed as follows. A chirped terahertz time domain spectroscopy scanned imaging system is used for making the image data measurements; such a system is illustrated in Fig. 10.2. The target is mounted on a motion stage so that the object can be rotated and linearly moved. As seen in Fig. 10.2, terahertz pulsed imaging (TPI) is achieved by repeating pulsed terahertz measurements in a 2D raster scan. The ultrafast pulsed laser beam is split into separate probe and pump beams. The path length of the pump beam is adjusted by a delay stage, then transmitted through a chopper and enters one of the optical rectification crystals, which operates as a terahertz emitter. A pair of parabolic mirrors are used to focus T-rays onto the sample. The T-rays emerging from the sample are adjusted again by another pair of parabolic mirrors, then combined with the probe beam co-linearly. For the current chirped pulsed imaging system, the optical probe beam is linearly chirped by a grating pair. The grating pair encodes the chirped signal in space and allows an electrooptic crystal to sample the terahertz temporal profile simultaneously. The grating pair (grating constant 10 μm) is setup so that the grating separation is 4 mm and the angle of incidence is 51°, giving a chirped probe pulse width of 21 ps. The probe pulse is linearly polarised by P1 and the polarisation modulation is converted to an amplitude modulation by polarizer P2 whose polarisation is perpendicular to P1. On transmission through the sample the THz radiation and optical probe beam are reflected by

NOTE:
This figure is included on page 207 of the print copy of the thesis held in the University of Adelaide Library.

Figure 10.2. A chirped probe pulse terahertz imaging system. It is used to realise terahertz CT imaging. In practice, the sample is placed on a rotational stage. After Ferguson *et al.* (2002b).

and transmitted through an indium tin oxide (ITO) THz mirror, separately. The detector crystal produces optical outputs that are proportional to the terahertz response. A CCD spectrometer is used for signal measurement. For the current reconstruction, we employ coherent spectrometry with a spectral resolution of 17 GHz and a detection lock-in amplifier time constant of 10 ms.

This current terahertz imaging setup achieves point to point detection. A linearly moving stage enables parallel scanning and a rotating stage allows projections of an object to be taken at a number of projection angles. Fig. 10.3(a) illustrates the terahertz sampling pattern. Adding to these two spatial variables is the time variable which is associated with the transient terahertz pulse. The measurements gathered from such a setup is organised as a three-dimensional (3D) data set, indexed by the variables (θ, ξ, t), where θ is the projection angle, ξ is the perpendicular distance from projection path to rotation axis, and t is the sampled time, as shown in Fig. 10.3(b). The optical properties of the material are extracted in the Fourier domain. In this domain, the tomographic inversion process takes place, to yield a final reconstruction as a function of the variables (x, y, ω), where ω indicates the frequency, and the x and y axes represent standard Cartesian coordinates. The processing required to effect this transformation is described in Section 10.2.

NOTE:
This figure is included on page 208 of the print copy of the thesis held in the University of Adelaide Library.

Figure 10.3. Schematic of THz CT sampling and the relative coordinate systems. (a) Illustration of data sample pattern of the current imaging. (b) Two sets of coordinate systems for the reconstructed terahertz CT. A typical terahertz pulse spectrum is shown as a function of time t in the inset. After Ferguson *et al.* (2002b).

10.3.1 Characteristics of diffraction grating pair

A conventional parallel grating pair is used to create a chirped pulse. A chirped pulse imaging system greatly accelerates the image acquisition process, because it allows one-shot acquisition removing the need for a delay stage. The different wavelength components of the incident pulse traverse different path lengths due to the variation in first order diffraction angle with wavelength.

According to Treacy (1969), for a parallel grating pair, if light of wavelength λ is incident on a grating with grating constant d at an angle γ (relative to the grating normal), and a wavelength-dependent angle θ is the angle between incident and diffracted rays, then the ray path length (ABCD in Fig. 10.4) is given by

$$p = b(1 + \cos\theta) = c\tau \tag{10.4}$$

where, $b = G\sec(\gamma - \theta)$ is a slant separation between the two gratings, if G is the perpendicular distance between the gratings, and τ denotes group delay.

The variation of group delay with wavelength is defined as

$$\delta\tau = \frac{b(\lambda/d)\delta\lambda}{cd[1-(\lambda/d-\sin\gamma)^2]}. \tag{10.5}$$

The group velocity dispersion (GVD) indicates the group delay dispersion per unit length, which represents the phenomenon of the frequency (or wavelength) dependent group velocity of a wave. Group velocity reflects the changes of the wave in propagated amplitude (known as the envelope of the wave). It is an important factor in controlling temporal pulse broadening.

The relation between γ and θ for first-order diffraction is

$$\sin(\gamma-\theta)+\sin\gamma = \frac{\lambda}{d} \tag{10.6}$$

here, d is the grating constant. Combined with Eq. (10.4), it is concluded that the optical path through grating pair is longer for the longer wavelengths than that of the shorter ones.

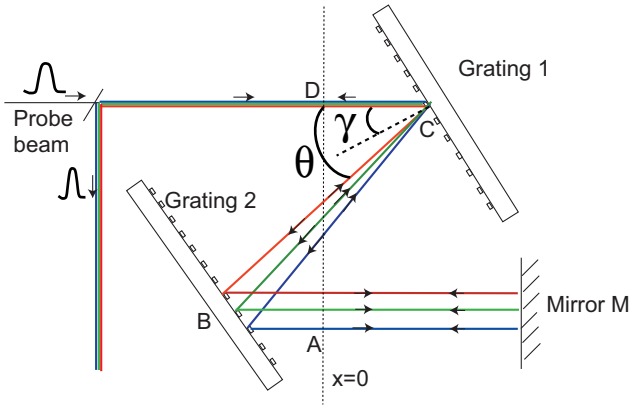


Figure 10.4. The geometry of a diffraction grating for chirped pulsed compression. The grating is used to impart a linear chirp to a laser pulse. The mirror M reflects the beam back into the grating pair. The optical path length is greater for longer wavelengths. The angle of incidence is γ and θ is the angle between incident and diffracted rays.

Fig. 10.4 shows a simplified pulse compressor via a pair of parallel gratings. The mirror M reflects the beam back into the grating pair, and the pair of parallel and face-to-face gratings is used to diffract light twice, so that the outgoing waves are parallel to, but laterally displaced from, the incoming waves. That is, pulse stretching is essentially the reverse of the pulse compression.

According to the grating equation, the longer wavelength light (red line) is diffracted through a smaller angle θ than the shorter wavelength light (green and blue line), so that the red light strikes the second grating to the left of the green while the blue light is to the right of the green. As a result, the total free-space path taken by a red ray is longer than that of a green ray, which is in turn longer than the blue ray's path. It turns out to be that the output from the grating is a pulse with a longer pulse duration and a wavelength that varies linearly with time. A negative group velocity is produced via the pair of parallel gratings. Since the input pulse is positively chirped and travels through the gratings, the output pulse becomes shorter due to partial cancellation of the positive GVD effect by the negative GVD of the gratings (Treacy 1969).

Electrooptic (EO) detection of a terahertz pulse using a chirped probe pulse was first demonstrated by Jiang and Zhang (1998b). This novel technique allows the full T-ray waveform to be measured simultaneously rather than requiring a stepped motion stage to scan the temporal profile. This provides a significant reduction in the acquisition time and greatly extends the applicability of T-ray systems in situations where the sample is dynamic or moving. Indeed, single shot measurements have been demonstrated for measuring a T-ray pulse using a single femtosecond light pulse (Jiang and Zhang 1998c).

However, this method degrades the SNR by spreading the available THz power over multiple pixels and diffraction effects can corrupt the temporal measurements. To avoid these additional concerns, the current experiment concentrates on the use of scanned imaging by focusing the THz pulses to a point and raster scanning the target.

10.4 Calculation of terahertz parameters for reconstruction of THz CT

One of the advantages that terahertz CT has over X-ray CT is that $s(\theta, \xi)$ may be one of several parameters derived from terahertz pulses. Fundamentally, a terahertz CT setup is capable of measuring the transmitted terahertz pulse as a function of time t, for a given projection angle and projection offset. In principle, terahertz sinograms can be obtained in both time and frequency domains.

10.4.1 Frequency domain sinogram for terahertz CT

The measured terahertz pulse is a function of time t, at a given projection angle and projection offset $p_d(t, \theta, \xi)$. Let us denote the Fourier transform of this time domain pulse by $P_d(\omega, \theta, \xi)$. The reference pulse $p_i(t)$ and the corresponding Fourier response $P_i(\omega)$ can be measured by removing the target object from background. If the target is rotated and probed by terahertz beams, $P_d(\omega, \theta, \xi)$ may be evaluated by adding sufficient projection angles to allow the filtered back projection algorithm to be applied at each specific frequency ω. This is based on the approximation that the detected terahertz signal is viewed as a linear integral of the incident terahertz pulse,

$$P_d(\omega, \theta, \xi) = P_i(\omega) \exp\left[\int_{L(\theta,\xi)} \frac{-i\omega\hat{n}(r)}{c} \mathrm{d}r\right] \tag{10.7}$$

where P_d and P_i are the Fourier transforms of the detected and incident terahertz signals, respectively; c is the speed of light in free space, L is the projection path, a straight line between the source and detector. The unknown complex refractive index of the sample is denoted by $\hat{n}(\omega, r) = n_\delta(\omega, r) + ik(\omega, r)$, where $n_\delta(\omega, r)$ is the real refractive index deviation and $k(\omega, r)$ is the extinction coefficient, related to absorption coefficient α via $k(\omega, r) = \alpha/2k_i$ (k_i is the incident extinction coefficient). Let us define that,

$$P_n \doteq \left[\frac{P_d(\theta, \xi)}{P_i(\theta, \xi)}\right] / k_i = \int_L n_\delta(r) dr = \Re\{n_\delta(r)\} \tag{10.8}$$

$$P_\alpha \doteq -2 \left\| \frac{P_d(\theta, \xi)}{P_i(\theta, \xi)} \right\| = \int_L \alpha(r) dr = \Re\{\alpha(r)\} \tag{10.9}$$

where $\arg(x)$ denotes the phase or argument of complex valued x, $\|x\|$ denotes the magnitude of the complex scalar x, and P_n and P_α are the projection data inputs to the filtered back projection algorithm as required to reconstruct n_δ and α, respectively, at a specific terahertz frequency ω. The sign r denotes the position of the incident field (the sensor). The frequency sinogram is applied to the vial and tube data sets (see Chapter 11) for THz image reconstruction experiments.

10.4.2 Time domain sinogram for terahertz CT

This method is based on the assumption that the target is dispersionless and therefore the THz pulse shape is unchanged after propagation through the target apart from attenuation and time delay. A reference terahertz pulse $p_r(t)$ is measured without the

target in place. To estimate the phase shift t of a terahertz pulse $p_s(t)$, the two signals are resampled at a higher rate using bandwidth-limited interpolation,

$$p_{\text{ref}}(m) = \sum_{t=-\infty}^{\infty} p_r(t)\text{sinc}\left[\frac{1}{q}(m - qt)\right] \tag{10.10}$$

$$p_{\text{sample}}(m) = \sum_{t=-\infty}^{\infty} p_s(t)\text{sinc}\left[\frac{1}{q}(m - qt)\right]. \tag{10.11}$$

The two interpolated signals are then cross-correlated, and the maximised cross-correlation product at each angle as the lag is taken as the estimation of the phase delay of $p_s(t)$. Mathematically, this process is described by:

$$\begin{aligned} R_{p_s p_r}[m] &= p_{\text{sample}}(m) \otimes p_{\text{ref}}(m) \\ &= \sum_{t=-\infty}^{\infty} p_{\text{sample}}(k)p_{\text{ref}}(k - m) \end{aligned}$$

$$T_d = \left\langle p_{\text{sample}} \otimes p_{\text{ref}}(m) \right\rangle_{\text{maxlag}} \tag{10.12}$$

where T_d is to estimate the delay time of terahertz phase, $p_{\text{sample}}(m)$ and $p_{\text{ref}}(m)$ are equal to $p_d(t)$ and $p_i(t)$ after interpolation by a factor of q. For the current experiment, $q = 2$ and $m = 301 \times 10 = 3010$. As shown in Eq. (10.12), the operator R denotes the cross-correlation and $\langle f(t)\rangle_{\text{maxlag}}$ denotes calculating the value of t at which the function f takes its maximum.

Timing sinogram can be calculated based on the following equation

$$p_{\text{time}} = \int_{L(\theta,\xi)} T_{\text{delay}} dr = \Re\{T_{\text{delay}}(r)\}, \tag{10.13}$$

here, p_{time} denotes the sinogram image in the time domain, recovered from the maximum time delay.

Fig. 10.5 illustrate the cross-correlation algorithm. The truncated terahertz projection response shown was interpolated and cross correlated with the reference pulse. The lag (T) at which the cross-correlation is maximised provides an accurate estimate of the delay between the two pulses. The algorithm is applied to cylinder data for terahertz CT reconstruction.

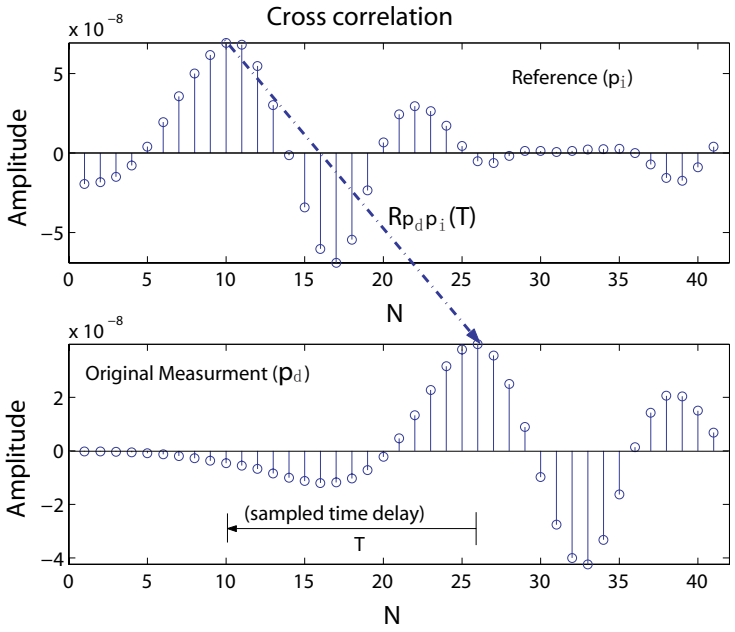


Figure 10.5. A cross-correlation algorithm. It is illustrated with the truncated projection of cylindrical data sets after interpolation for clarity.

10.5 Chapter summary

This Chapter introduces computed tomography applied to terahertz imaging. We present the difference between the conventional Radon transform reconstruction and our modified Radon transform for terahertz computed tomography. It is an important basis for the research with respect to THz segmentation application on a volume space, presented in Chapter 11, and for wavelet based local tomography via pulsed THz detection, and continuous wave THz reconstruction with a QCL, presented in Chapter 12 and Chapter 13, respectively. In next Chapter, the wavelet scale correlation segmentation technique is adopted to achieve material discrimination of THz measurements. It turns out that only a segmenter parameter needs to be adjusted.

Chapter 11

2D Wavelet Segmentation in 3D T-ray CT

IN this Chapter, segmentation techniques for terahertz (T-ray) computed tomographic (CT) imaging are investigated. A set of linear image fusion and novel wavelet scale correlation segmentation techniques is adopted to achieve material discrimination within a three dimensional (3D) object. The methods are applied to a T-ray CT image data set taken from a plastic vial containing a plastic tube. This setup simulates the imaging of a simple nested organic structure, which provides an indication of the potential for using T-ray CT imaging to achieve T-ray pulsed signal classification of heterogeneous layers. The wavelet based fusion scheme enjoys the additional benefit that it does not require the calculation of a single threshold and there is a single parameter to adjust.

11.1 An introduction for THz segmentation experiments

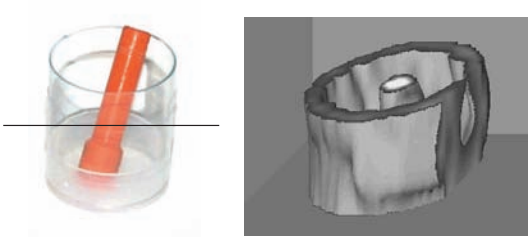
Current two- and three- dimensional T-ray CT imaging is affected by different types of noise, which can restrict its usefulness. These drawbacks have prevented successful development of accurate and reliable edge detection and segmentation algorithms for T-ray CT images. A reliable and automatic segmentation detection is of great importance in many applications of T-ray image processing, especially for the T-ray pulsed signal classification of 3D objective CT imaging with heterogeneous layers. Handley *et al.* (2002) employed clustering techniques for segmentation with wavelet preprocessing, on real and synthetic THz CT slices.

The current wavelet scale correlation based segmentation-by-fusion algorithms afford another approach in 3D T-ray tomographic imaging to probe the interior regions of objects. In conjunction with various edge detection techniques, our algorithm provides quantitative information for measurement and classification applications. One of the main advantages of the algorithm is that, in its simplest form, it requires only a single segmentation parameter. In addition, a wavelet based segment-by-fusion detection method for T-ray images is effective with respect to noise.

This Chapter contains two main Sections: methodology and experimental results. The methodology Section describes the algorithms employed in this work. The chirped probe pulse THz imaging system, described in Section 10.3, reviews the 3D T-ray CT system that is used to acquire the data, after Ferguson *et al.* (2002b), used for our signal processing experiments. CT reconstruction algorithms are discussed in Section 10.2. In Section 11.3 a novel method is introduced to achieve reconstructed segmentation, called wavelet scale correlation based segmentation-by-fusion. Section 11.4 consists of the experimental segmentation results and the evaluation of segmentation quality. Section 11.5 concludes this Chapter.

11.2 Representation of a target sample

The object photograph of a plastic vial containing a different plastic tube, simulates a simple nested structure—as shown in Fig. 11.1(a). The target is imaged with a 1 mm step size in the x and y dimensions, and at projections separated by 10^{o}. First the reconstruction is performed using the timing of the peak of the THz pulse in the time domain to yield a reconstruction of the bulk absorption coefficients. For the current experiment,



(a) (b)

Figure 11.1. A nested structure for the illustration of T-ray CT system. A vial containing a plastic tube is used as a simple nested structure to test the T-ray CT system and to demonstrate the resulting reconstruction and classification of the T-ray CT imaging. (a) The optical image of the vial and the tube. (b) The 3D reconstruction of the vial and tube based on the T-ray CT slices.

the real part, instead of using imaginary part, of the expression $\log[P_d(\omega)/P_i(\omega)]$ is used to reconstruct T-ray CT at each sampled frequency, as the former shows the ability to reconstruct a relatively accurate and clear tomographic image. The central slice is reconstructed at each of the ten lowest frequencies, from 0.0213 to 0.213 THz, (illustrated in Fig. 11.2). The 3D rendered image is illustrated in Fig. 11.1(b), which is the combination of the reconstructed slices at a number of target heights. The resulting isosurface is constructed using the pixels where the reconstructed absorption coefficients are evaluated. The reconstructed image dimensions are rather accurate, with the vial and cylinder diameters being within 15% of the actual dimensions measured with calipers. However, the vial thickness is much thicker than expected because of the coarse reconstruction grid size of 1.5 mm. The grid size may be improved using more projection angles at smaller angular intervals, but also needs improved 'Δx' resolution, i.e. longer sinograms.

11.3 Wavelet based segmentation by fusion

As described above, T-ray CT extracts the frequency dependent 3D characteristics of a target sample. This makes possible rich, four dimensional data sets that describe the sample. To achieve the final T-ray CT image, an inverse Radon transform (IRT) is computed on the Fourier coefficients of the measured signals. In order to obtain the material discrimination within a three dimensional (3D) object, a set of linear image

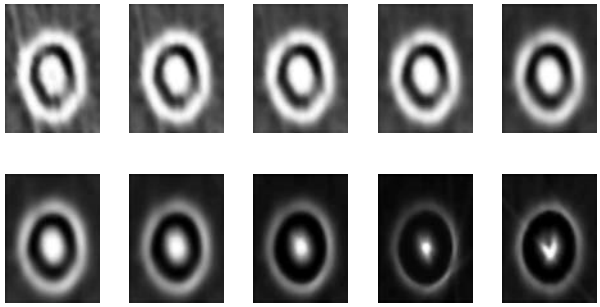
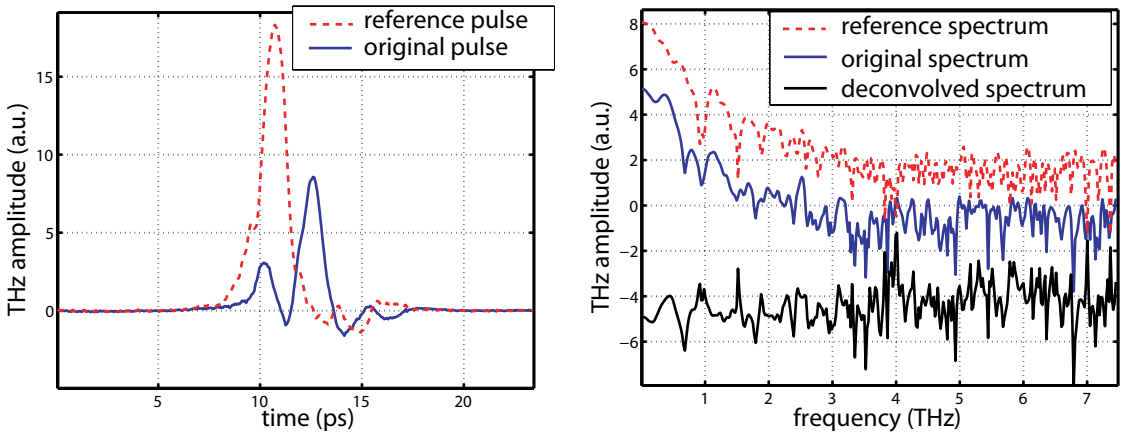


Figure 11.2. Illustration of the reconstructed T-ray CT slices. These slices are reconstructed at the first 10 frequencies, in increasing order from top left at the object height of 7 mm.



(a) (b)

Figure 11.3. Illustration of T-ray signals and spectra. (a) Detected T-ray signal and reference; (b) their spectra in log scale with offsets of 2 a.u. and 4 a.u. corresponding to the original and de-convolved spectra, respectively, for clarity.

fusion and novel two dimensional (2D) wavelet scale correlation segmentation algorithms are adopted. The methods are applied to a T-ray CT image data set of a plastic vial containing a plastic tube, with an aim to image a simple nested structure. The setup is imaged at various heights, ranging from 5 mm to 9 mm (from the bottom), in 1 mm increments. Image fusion algorithms are conducted on the low frequency reconstructed T-ray CT images at different target heights for a smooth version of the CT images. In turn, a 2D discrete wavelet transform (DWT) is taken on the fused image. In order to reconstruct the same size of the approximate sub-images as the fused image, the inverse discrete wavelet transform (IDWT) is performed after the appropriate

subband is isolated. The target segments are localised by computing the correlation between the subimages corresponding to the first two approximate scales.

11.3.1 Image fusion of T-ray CT images for a 3D target

A fused image can be created via merging two or more images, acquired from a single source to obtain the sharp display of various materials and structures (Brown 1992). In this case, much more meaningful visual information can be achieved in the combined images than from the individual ones.

Our segmentation techniques are based on a plane fusion approach (Jan 2006) to match and fuse multiple desirable slice images at various frequencies into a single slice image. A linear combination of weighted slice images is used to perform plane image fusion in the current experiment. The weights are selected based on the various intensities for the different frequency of T-ray CT images. In this work, the weighting scheme is empirically chosen to be 1/15 for the five lower frequencies and 2/15 for the five higher frequencies to compensate for the attenuation in signal strength in the higher frequency components. The resultant fused image has clear target contours and shows strong contrast between the target regions and the background. Fig. 11.4(a) shows the resultant fused image with image size of 89×89 pixels.

11.3.2 Discrete wavelet transforms in two dimensions

The 2D discrete wavelet transforms (DWT) can be realised by digital filters and down-sampling the T-ray image. Expansion in a 2D scaling function, $\phi(x,y)$ and three 2D wavelets functions, $\psi^H(x,y), \psi^V(x,y)$, and $\psi^D(x,y)$, are calculated by taking the 1D fast wavelet transform (FWT) of the rows of an image and the resulting columns (Mallat 1999). This algorithm can be inverted via up-sampling the signals, filtering and adding the output together to recover the low-frequency subband at the previous level of the multiresolution analysis. The process is repeated for the depth of the DWT to obtain the original image. In our work, however, the individual subbands at the lowest resolution are isolated (i.e. all other subbands at the same level are set to zero) before application of the inverse wavelet transform (IWT). This processing allows the wavelet scale correlation based segmentation to be performed.

Fig. 11.5(a) and (b) are the reconstructed sub-images after taking 2D inverse wavelet transform of approximate coefficients at the first and second wavelet decomposed

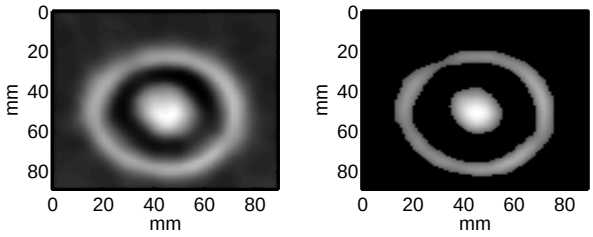


Figure 11.4. (a) Fused image via merging T-ray CTs at ten lowest frequencies. (b) Final extracted object segments from the background using wavelet based segmentation by fusion.

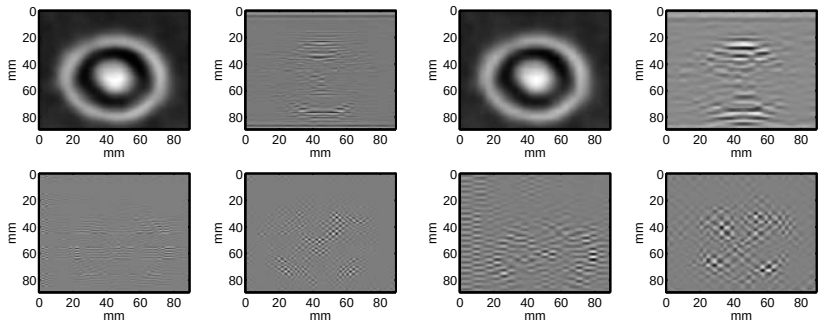


Figure 11.5. (a) The reconstructed approximate sub-image after one-scale wavelet decomposition of the fused T-ray CT. (b) The reconstructed approximate sub-image after two-scale wavelet decomposition of the fused T-ray CT.

scales. The T-ray fused image at a 7 mm target height is used. A Daubechies 4 wavelet transform is applied for this case. In each group of the two scales of reconstructed T-ray CT sub-images, the first sub-image is the reconstructed approximate sub-image and the other three are detail reconstructed sub-images. It can be seen that the approximate sub-images are smoothed versions of original image.

11.3.3 2D wavelet scale correlation based segmentation

The aim of segment detection is to achieve differentiated subdivision of constituent regions of an image. The method used in this Thesis is motivated by one dimensional wavelet scale correlation denoising. Firstly, the target cross-sections are assumed to be corrupted by additive white Gaussian noise, which is randomly distributed. The target objects are separated by their absorption coefficients, which are indicated by

the intensity in the images. With the incremental wavelet scale, the noise is reduced and the target intensity (energy) is increased in an image. After extracting the low frequency 2D subband of fused T-ray CT images, an increased energy with an increase in wavelet scale is used as a cue to extract the target regions. This cue is consistent with the procedure used in 1D wavelet de-noising, where the noise usually appears as small coefficients in the wavelet subbands. By computing a correlation of two scales of wavelet subbands, the signal components should survive in large coefficients while the noise would be diminished as it would be spread over many small coefficients. The wavelet scale correlation based segmentation algorithm is summarized as:

1. Calculate 2D reconstructed approximation images at first scale $S_1(m,n)$ and second scales $S_2(m,n)$, i.e. for transform depths of 1 and 2, respectively.

2. Compute the correlation $R_{1,2}$(m,n) for the two scales:

$$R_{1,2}(m,n) = S_1(m,n) \times S_2(m,n). \tag{11.1}$$

3. Compute the energy of $R_{1,2}(m,n)$ and $S_1(m,n)$:

$$E_{S1} = \sum S_1^2(m,n) \tag{11.2}$$

$$E_{R1,2} = \sum R_{1,2}^2(m,n). \tag{11.3}$$

In order to make it comparable between the wavelet coefficients and the correlation coefficients, it is necessary to normalise the coefficients:

$$R_{1,2}^*(m,n) = R_{1,2}(m,n)\sqrt{E_{S1}/E_{R1,2}}. \tag{11.4}$$

4. The wavelet energy $|R_{1,2}^*(m,n)|$ and $\lambda|S_1(m,n)|$ is compared. If $|R_{1,2}^*(m,n)| > \lambda|S_1(m,n)|$, ($\lambda$ is a parameter, chosen to be 1 for this experiment), the pixel at (m,n) is extracted as part of a target segment, otherwise it is regarded as background.

The Canny edge detector, combined with Otsu's threshold method (Gonzalez and Woods 2002), is used to perform the final subtraction of each target function edge in the T-ray CT image. It should be noted that the current algorithm only requires the adjustment of a single parameter λ in the processing.

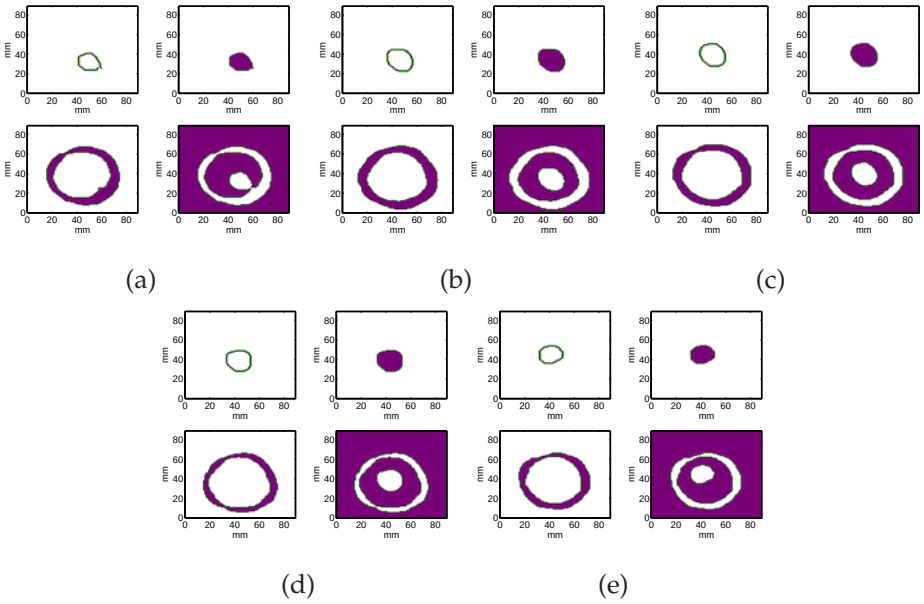


Figure 11.6. Segmented images at different heights. These segmented image sets at heights of (a) 5 mm; (b) 6 mm; (c) 7 mm; (d) 8 mm and (e) 9 mm from the bottom of the vial. The four subfigures, clockwise from top left, illustrate the tube edges, tube segments, the air segments (background) and vial segments, respectively.

11.4 Experimental result

11.4.1 Extracted object segments

The algorithm described in Subsection 11.3.3 is applied to extract segments corresponding to the vial and tube. The extracted approximate sub-images, after computation of the 2D DWT of the 89×89 fused images, are shown in Fig. 11.5(a) and (b); the Daubechies 4 wavelet is used in this case, as mentioned above. The result of the correlation is shown in Fig. 11.4(a). Figures 11.6(a)-(e) illustrate the resultant segment subimages corresponding to heights of 5 mm to 9 mm from bottom to top. The upper left portion in each subfigure is the edge subimage of the tube; the upper right portion is the segment subimage of the tube; the lower left portion is the vial segment region; the lower right portion is the air segment region.

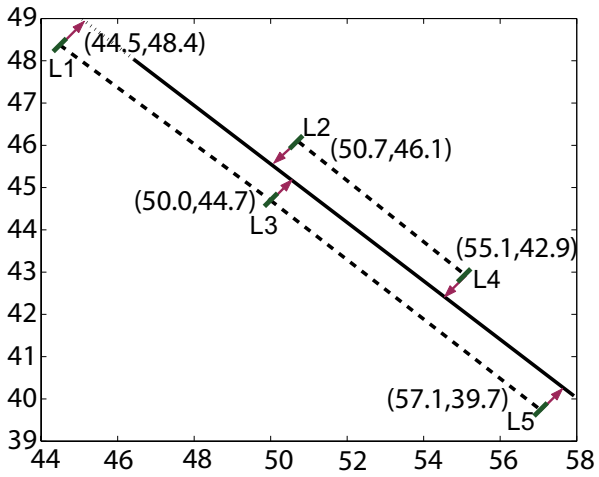


Figure 11.7. Illustration of the centroid locations of the target tube segment. The centroids correspond to five target heights, L1 to L5. The three dotted lines are the connections of each of the two measured centroids from five target heights; the solid line is to fit the five measured centroids. The X and Y axes represent the position of each pixel with the unites of X and Y in 'pixel'.

11.4.2 Segmentation quality

The segmentation quality cannot be directly verified since the ground truth is not known. This situation is typical of tomographic applications, where it is often not possible to dissect the subjects to provide knowledge of the ground truth. Alternatively, a comparison with manual segmentation is undesirable, since that depends on the selected threshold, which cannot afford a standard resultant segment for comparison.

In this experiment, we exploit the fact that the internal structure, the tube, is straight. This implies that the segmented plastic tube positions (see top right of each sub-figure in Fig. 11.6) should be proportionally displaced from each other, since they correspond to constant increments of 1 mm in height. To obtain this measurement, we find the centroid of the extracted vial tube for each height (labelled L1-L5) and the resultant x and y locations are plotted in Fig. 11.7. The achieved linear regression line, with slope of -0.7 and an offset of 81.54 pixels, is used to fit the five measured centroids. The mean square error of each point to the given solid line is 1.034. This indicates that our algorithm is capable of locating the centroid of the plastic tube to within a handful of pixels. Given the number of noise sources and the quality of the reconstructed images, this error may be considered to be reasonable.

11.5 Chapter summary

This Chapter investigates an application of wavelet based segmentation by fusion. It is a first exploration of automatic T-ray CT target *identification* within 3D heterogeneous structures. The algorithm successfully segments different target regions and is able to correctly locate the regions to within a few pixels. This provides motivation for future application of classification algorithms for material identification of 3D heterogeneous layers. For future work, it is suggested that the algorithm be characterised further by studying the effect of the segmentation parameter λ on the results. In addition, more experimental work is needed to verify the general performance of this algorithm.

A further application of wavelet transform scheme to achieve wavelet based local tomography via broadband pulsed THz sources is experimentally illustrated in the next Chapter, which shows how important wavelet transforms are for the achievement of THz computed tomographic reconstruction, especially in the region of interest.

Chapter 12

Wavelet-Based Terahertz Coherent Local Tomography

IN the previous Chapter we explored wavelet-based coherent tomography and in this Chapter we now build on this approach to specifically perform *local* tomography. With traditional CT techniques such as X-ray tomography, full exposure data are needed for inverting the Radon transform to produce cross sectional images. This remains true even if the region of interest is a small subset of an entire image. For time-domain terahertz measurements, the requirement for full exposure data is impractical due to the slow measurement process. In this Thesis, we apply a wavelet-based algorithm to reconstruct THz CT images with a significant reduction in the required measurements when the region of interest is small. Instead of inverting the Radon transform, the approach uses modified wavelet and scaling ramp filters along with the traditional back projection algorithm to obtain the resultant reconstruction.

The algorithm recovers an approximation of the region of interest (ROI) from terahertz measurements within its vicinity, and thus improves the feasibility of using terahertz imaging to detect defects in solid materials and diagnose disease states for clinical practice, to name a few applications. This novel approach is illustrated in two terahertz imaging case studies.

12.1 Motivation

The aim of our current work is to perform terahertz tomographic reconstructions, particularly localised reconstructions based on limited data. A curious and unfortunate aspect of CT in two-dimensions is that the reconstruction of a small portion of an object via traditional filtered back projection (FBP) requires a complete set of projection data, which leads to increased measurements. Mathematically, the problem with the conventional FBP algorithm for local reconstruction is that the Hilbert transform in the inversion process de-localises the local measurements (Holschneider 1991, Rashid-Farrokhi *et al.* 1997). In traditional X-ray CT, the implication of this fact results in greatly increased exposure of patients to harmful radiation. Within the context of terahertz CT, the time required to make complete measurements can render the technique impractical, or even inaccurate, due to measurement errors running away with laser drift, for instance.

In order to overcome this problem, the Soviet scientists Vainberg *et al.* (1981) invented Λ-tomography in the 1980s. Strictly speaking, Λ-tomography is not true local tomography, for it only provides a transformed version of the underlying distribution. In the pursuit of true local tomography, a wavelet technique was first proposed for the assistance of mapping arbitrary distributions by Holschneider (1991), who laid the mathematical foundation for subsequent work on wavelet localised tomography. In 1994, Olson and DeStefano (1994) carried out a one dimensional wavelet transform on the sparsely sampled projections, and then the standard FBP algorithm was used to recover a region of interest (ROI) via a group of detailed coefficients and another group of approximate coefficients for the remaining image. Similar to the method adopted by DeStefano and Olson, Delaney and Bresler used a two-dimensional (2D) wavelet transform in the projection domain and FBP was utilized for reconstruction of the local image from the wavelet-filtered projections (Delaney and Bresler 1995). Neither of these approaches are genuinely localised, for they both relied on the measurements from outside of the ROI. A truly local CT reconstruction using wavelet is introduced by Rashid-Farrokhi *et al.* (1997), where 2D discrete wavelet ramp filters are applied to the local projection and resulted in the local reconstruction of both high- and low- frequency features of an image. Later, Zhao (1999) suggest an efficient subband coding algorithm for wavelet based filtered back projection; and Madych (1999) reconstruct local images via simple modification of traditional FBP algorithms and especially emphasize

that continuous wavelet transforms facilitate good performance for high-resolution local reconstruction. More recently, a couple of novel methods related to wavelet based local reconstruction are conducted via investigating a set of modified one-dimensional projections as a function of the space variable by Smith and Adhami (2000) and via suggesting a well chosen analysis 'wavelet' by Bilgot *et al.* (2004).

The main goal of this Section is to present a wavelet based reconstruction algorithm for terahertz computed tomography and 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 reconstruction. The current algorithm provides new insight into the relationship between local reconstruction, local projection, and the resolution of terahertz coherent tomography.

The current algorithm achieves the reconstructed approximate and detailed portions of an image via conducting back projection on ramp filtered scaling and wavelet functions. Compared to previous algorithms in THz CT (Ferguson *et al.* 2002a, Nguyen *et al.* 2006), the current reconstruction algorithm accelerates terahertz image scanning and reduced computation complexity, as only a small number of projections on lines passing close to the ROI are required. Therefore, it is computationally more efficient. Uniform exposure is adopted at all angles for simpler implementation in the hardware setup. Reconstruction ability at off-centered and centered regions of interest are also explored.

This Chapter consists of several Sections. This imaging experiment is based on a terahertz functional imaging system. Based on the basics of the wavelet transform, mentioned in Chapter 6, a full-data reconstruction technique based on the wavelet transform is also involved, represented in Section 12.2. Following this, local tomographic reconstruction algorithms via wavelet transforms are introduced in Section 12.3. Then Section 12.4 discusses the implementation of this method, and in Section 12.5, the tomographic results are presented.

12.2 Two dimensional wavelet based CT reconstruction

This Section briefly describes an algorithm, which is applied to obtain the wavelet coefficients of a function on $\mathbb{R}^2$ space, based on Radon transform data. This method enables reduced computation compared to the wavelet coefficients obtained, after conducting wavelet transforms in a reconstructed image. Moreover, the wavelet coefficients are

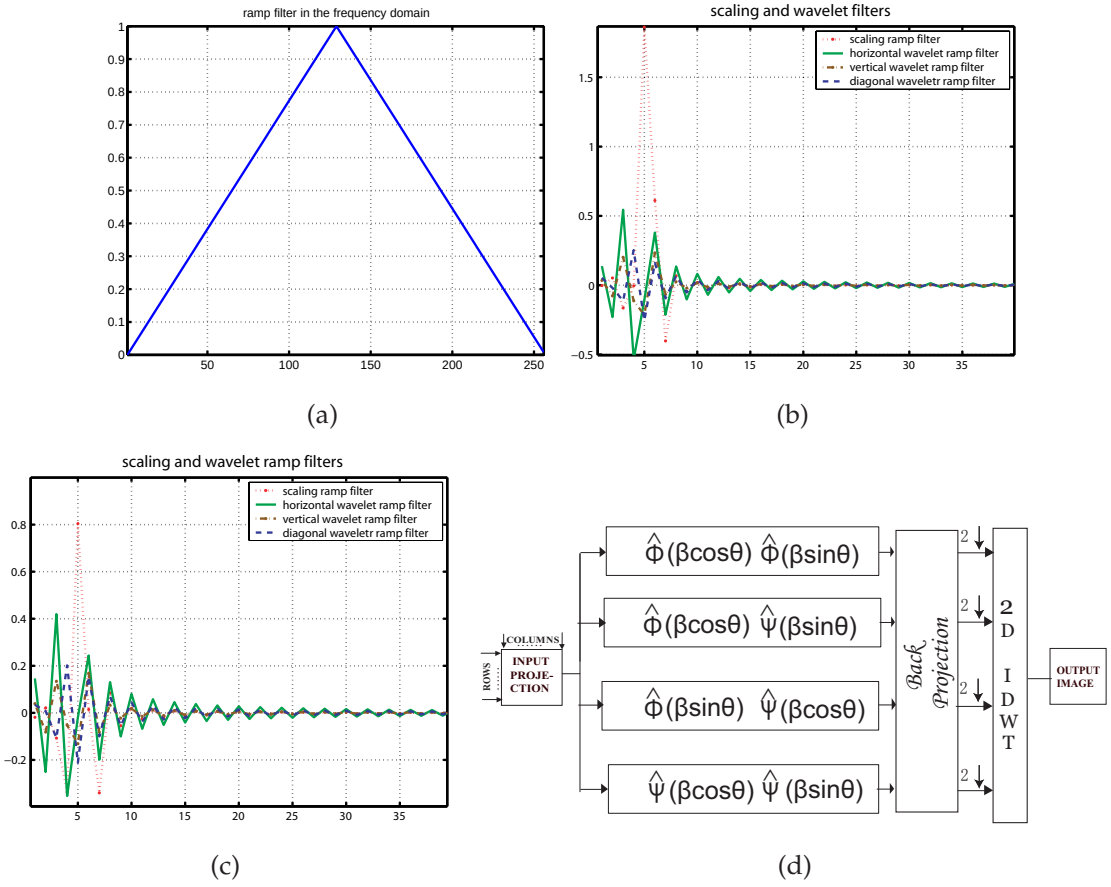


Figure 12.1. Wavelet ramp filters. (a) Illustration of a traditional ramp filter. (b) and (c) Illustrate the scaling and wavelet ramp filters at the sixth projection angle (43.2 degree) using BiorSplines 2.2 wavelet, respectively, (d) Block diagram of wavelet reconstruction from projection data. The 2D inversion of the traditional wavelet transform (2D IDWT) is conducted on the back projection of the approximate and detail sinograms, after downsampling by a factor of 2 (indicated by down arrow) is performed.

calculated locally allowing the local reconstruction to yield local computed tomography (Rashid-Farrokhi *et al.* 1997). The main formulae for 2D DWT, on projection data, for the reconstruction of a CT image are introduced, which are implemented via performing separate wavelet transforms on 1D projection data.

The filtered back projection algorithm for terahertz CT reconstruction is expressed as follows:

$$I(x,y) = \int_0^{\pi} \mathrm{d}\theta \cdot \left[\int_{-\infty}^{\infty} S(\theta,\beta)|\beta| G_{2^j}(\beta\cos\theta, \beta\sin\theta) \exp(i2\pi\beta\xi)\mathrm{d}\beta \right] \tag{12.1}$$

where $S(\theta,\beta)$ and $G_{2^j}(\beta_1,\beta_2)$ are the spatial Fourier transforms of $s(\theta,\xi)$ and g_{2^j} (a wavelet ramp filter in the time domain), respectively.

The function enables image reconstruction as the conventional inversion of the Radon transform method, while the ramp filter $|\beta|$ is replaced by the wavelet ramp filter $|\beta|G_{2^j}(\beta\cos\theta,\beta\sin\theta)$.

As for a separable wavelet basis, the approximate and detail coefficients have been given in Eq. (6.49). We rewrite it here

$$\begin{cases} c_j(k,l) = \langle f(x,y),\phi_j(k,l)\rangle \\ d_j^H(k,l) = \langle f(x,y),\psi_j^H(k,l)\rangle \\ d_j^V(k,l) = \langle f(x,y),\psi_j^V(k,l)\rangle \\ d_j^D(k,l) = \langle f(x,y),\psi_j^D(k,l)\rangle \end{cases} \tag{12.2}$$

where

$$\phi_j(x,y) = \phi_j(x)\phi_j(y) \tag{12.3}$$
$$\psi_j^H(x,y) = \phi_j(x)\psi_j(y) \tag{12.4}$$
$$\psi_j^V(x,y) = \psi_j(x)\phi_j(y) \tag{12.5}$$
$$\psi_j^D(x,y) = \psi_j(x)\psi_j(y) \tag{12.6}$$

where ϕ and ψ are the one-dimensional scaling and wavelet functions, respectively.

These coefficients can be calculated from the projection data via the first integral item in Eq. (12.1), replacing $G_j(a,b)$ by $\Phi_j(a,b)=\hat{\phi}_j(a)\hat{\phi}_j(b)$, $\Psi_j^H(a,b)=\hat{\phi}_j(a)\hat{\psi}_j(b)$, $\Psi_j^V(a,b)=\hat{\psi}_{2^j}(a)\hat{\phi}_j(b)$, and $\Psi_j^D(a,b)=\hat{\psi}_j(a)\hat{\psi}_j(b)$, respectively, where $a=\beta\cos\theta, b=\beta\sin\theta$, and the caret sign indicates Fourier transforms of the relative functions. For instance, the approximate coefficients are obtained by

$$c_j(k,l) = \int_0^{\pi} \mathrm{d}\theta \cdot \left[\int_{-\infty}^{\infty} S(\theta,\beta)|\beta|\Phi_j(\beta\cos\theta,\beta\sin\theta)\exp(i2\pi\beta\xi)\mathrm{d}\beta\right]. \tag{12.7}$$

The detail coefficients can be found in a similar way as follows:

$$d_j^i(k,l) = \int_0^{\pi} \mathrm{d}\theta \cdot \left[\int_{-\infty}^{\infty} S(\theta,\beta)|\beta|\Psi_j^i(\beta\cos\theta,\beta\sin\theta)\exp(i2\pi\beta\xi)\mathrm{d}\beta\right]. \tag{12.8}$$

This means that the filtered back projection method can be applied to the wavelet transform of the projections for image recovery, while the ramp filter is replaced by,

$$\begin{cases} H_\theta^C = |\beta|\Phi_j(\beta\cos\theta, \beta\sin\theta) = |\beta|\Phi_j(\beta\cos\theta)\Phi_j(\beta\sin\theta) \\ H_\theta^{D^H} = |\beta|\Psi_j^h(\beta\cos\theta, \beta\sin\theta) = |\beta|\Phi_j(\beta\cos\theta)\Psi_j(\beta\sin\theta) \\ H_\theta^{D^V} = |\beta|\Phi_j^V(\beta\cos\theta, \beta\sin\theta) = |\beta|\Psi_j(\beta\cos\theta)\Phi_j(\beta\sin\theta) \\ H_\theta^{D^D} = |\beta|\Psi_j^d(\beta\cos\theta, \beta\sin\theta) = |\beta|\Psi_j(\beta\cos\theta)\Psi_j(\beta\sin\theta) \end{cases} \tag{12.9}$$

where H_θ^C and $H_\theta^{D^i}, (i = H, V, D)$ are called scaling and wavelet ramp filters. Fig. (12.1) (b) and (c) illustrate the scaling and wavelet ramp filters at the sixth projection angle (43.2 degrees) using the BioSplines2.2 biorthogonal wavelet.

For the current image reconstruction, only one 2D wavelet transform step is used. This is because the single level decomposition of scaling and wavelet ramp filters allows clear reconstruction of an image in the ROI and it avoids more computational complexity due to more levels of WT employed (Delaney and Bresler 1995, Rashid-Farrokhi *et al.* 1997). The wavelet reconstruction formulae in Eq. (12.1) allow for such reconstruct by setting $j = 1$. Fig. (12.1)(d) shows the block diagram of the wavelet based reconstruction system. The 2D inversion of the traditional wavelet transform (IWT) is conducted after the back projection of reconstructed approximate and detail sinograms.

12.3 Local reconstruction using wavelets

A characteristic of many wavelets is a large number of vanishing moments. Hilbert transforms of functions with many vanishing moments have been shown to decay very rapidly at infinity (Delaney and Bresler 1995). In other words, a wavelet function with compact support allows a local basis to maintain its localised features after Hilbert transformation (Delaney and Bresler 1995, Berenstein and Walnut 1994). Fig. (12.1) (a)-(c) illustrates the ramp filter over the full frequency domain, the BiorSplines biorthogonal scaling and wavelet filters and the ramp filtered version of the BiorSplines biorthogonal wavelet and scaling filters, where the x axis means the number of time or frequency samples, and y axis means the relative amplitude. Fig. (12.1)(c) essentially shows the essentially compact support after applying Hilbert transforms. Therefore, the wavelet and scaling coefficients for some wavelet basis can be calculated after applying the projections passing through the region of interest plus a margin for

the support of the wavelet and scaling ramp filters. These reconstructed coefficients, in this experiment, are then directly applied to the inverse wavelet transforms for terahertz image reconstruction.

12.3.1 Error analysis

As the Hilbert transform is nonlocal, in order to recover an image in the local region of interest, the nonlocal data that is outside the ROI have to be considered for an accurate local reconstruction. An upper bound for the reconstruction error, denoted by ϵ_{ub} can be calculated and therefore the relative error ϵ_{rel} is used in this Thesis for evaluation of the current reconstruction. The principle for error analysis (Rashid-Farrokhi *et al.* 1997) is to consider the measured projections divided into two parts: one is for the region of exposure and another is its complement. We consider the original measurement of target image $I(x,y)$ with a support centered at the origin and within a radius of 1 unit. Based on the Cauchy-Schwartz inequality and the assumption: $|s_{\theta_k}| \leq 2\max|I(x,y)|$, where the support of $I(x,y)$ is within a radius of 1, $|s_{\theta_k}|$ is the projection at the k^{th} projection angle, the relative error ϵ_{rel} can be calculated via dividing the upper bound of the reconstructed error ϵ_{ub} by the maximum intensity of the support $\max|I(x,y)|$. The relative error calculated in the frequency domain satisfies the following equation:

$$I(x,y) = \frac{2\sqrt{2}\pi}{k}\frac{\sqrt{R-r_e}}{R}\sum_{k=1}^{K}\left(\sum_{|n|=-R}^{R}|[\bar{H}_{\theta_k}(n) - \bar{H}^{\text{T}}_{\theta_k}(n)]|^2\right)^{1/2} \tag{12.10}$$

where $\bar{H}_{\theta_k}$ is the inverse Fourier transform of H_{θ_k} and H_{θ_k} can be replaced by the ramp filter $|\beta|$.

The truncated filter $H^{\text{T}}_{\theta_k}$ is defined as

$$H^{\text{T}}_{\theta_k}(n) = \begin{cases} H_{\theta_k}(n) & \text{if } |n| < r_e - r_i, \\ 0 & \text{otherwise}, \end{cases} \tag{12.11}$$

here, r_e and r_i are the radii corresponding to ROE and ROI, both of which are centered at the origin of the image.

The upper bound of the relative error in the reconstructed sinogram of scaling and wavelet coefficients can be obtained by replacing H_{θ_k} in Eq. (12.10) with in Eq. (12.9) and multiplying by a normalizing factor. Eq. (12.12) represents the relative error in the reconstructed image from approximate reconstruction coefficients,

$$|\epsilon_{\text{rel}}(x,y)| = N_{I_{g,l_C}} \cdot \frac{2\sqrt{2}\pi}{k} \frac{\sqrt{R-r_e}}{R} \sum_{k=1}^{K} \left(\sum_{|n|=-R}^{R} |[\bar{H}_{\theta_k}(n) - \bar{H}^{\text{T}}_{\theta_k}(n)]|^2 \right)^{1/2} \tag{12.12}$$

where $N_{I_{g,l_C}}$ is the normalised scale factor of an image in relation to approximate reconstruction coefficients, which is calculated via dividing maximum intensity of global reconstruction, denoted by $\{I_r(x,y)\}$, by maximum intensity of local reconstruction regarding approximate wavelet coefficients, denoted by $\{I^C_{\text{local}}(x,y)\}$. The scale factor is defined as

$$N_{I_{g,l_C}} = \max |I_r(x,y)| / \max |I^C_{\text{local}}(x,y)|. \tag{12.13}$$

In the current experiment, for convenience, the error calculation is limited to the error of the approximate image reconstruction. For the calculation of relative error using the whole version of the wavelet reconstructed image, please refer to Rashid-Farrokhi *et al.* (1997) or Appendix C.

12.4 Implementation

12.4.1 Experiments

The current research based on terahertz imaging is most closely related to Rashid-Farrokhi *et al.* (1997). In this work, we experiment with the 2D wavelet technique using terahertz tomographic data by modifying the measured projections. As we show later, this modification involves an extrapolation technique to avoid edge effects due to sinogram truncation. It is observed that approximate coefficients of a scaling function show good localised features in the local reconstruction using our algorithm, where the reconstructed intensity of an image varies much between different target materials. It should be noted that, in the application of terahertz data for local reconstruction, it is found that the intensity at the edges of the region of exposure (ROE), where nonlocal

NOTE:
This figure is included on page 233 of the print copy of
the thesis held in the University of Adelaide Library.

Figure 12.2. Target object photographs. (a) An optical image of a target with 2 mm diameter holes drilled into a polystyrene cylinder with varying interhole distances. (b) Target object photograph with simple nested structure. The line indicates the measurement height of 7 mm. After Ferguson *et al.* (2002b).

data are set to zero, varies considerably after conducting either a traditional ramp filter or scaling and wavelet ramp filters.

In local reconstruction, artifacts are common close to the boundary of the ROE, which can readily be observed in the application to terahertz CT data. It is possible that reconstruction after applying a constant linear extrapolation results in missing information. This situation is illustrated in Fig. 12.5. In this experiment, there are two sets of terahertz data considered for reconstruction: a cylinder with holes inside—see the target photo in Fig. 12.2(a)—and a nested structure of a tube inside a vial. For the first set of terahertz data (the sample photo in Fig. 12.2(b)), with 101 projections at each of 25 projection angles covering a 180° projection area in a 100×100 pixel image. The line in the photo indicates the measurement height of 7 mm. Two situations are analyzed for this target sample: (i) an ROE of diameter 42 pixels at the center of the image and (ii) an ROE of diameter 67 pixels offcenter to the image. For the second set of terahertz measurements, with 51 projections at each of 36 projection angles covering a 360° projection area in a 100×100 pixel image, an ROE of diameter 18 pixels at the center of the image is explored. Each of data set has a pixel interval of 0.5 mm.

In order to recover the cross-sectional image in the region of interest, the values of the sinograms outside of the ROE are set to zero. The traditional filtered back projection formulae, Eq. (10.2), and wavelet based reconstruction, Eq. (12.1), are applied to the remaining projections, respectively for analysis and comparison. The original terahertz sinogram image for the first set of terahertz data can be calculated via applying Eq. (10.9).

Case study ♯ 1: Polystyrene target

We consider reconstruction at the center of the image, based on the first set of data. A cross correlation between the reference and signal is used to get the time-domain measurement aiming for the calculation of Radon transforms of terahertz data. Fig. 12.3(a) and (b) shows the original calculated projection and the projection with zero outside of ROE at the 25^{th} projection angle. Fig. 12.3(c) shows sharp variations along the borders of the ROE after applying wavelet ramp filters and ramp filter, respectively, on each of the 1D projections, which result in an image appearing relatively weakened in intensity when compared to a large constant bias that exists along the reconstructed edges in the region of interest. The constant extrapolation we use is given by Eq. (25) in Rashid-Farrokhi *et al.* (1997). In order to fit our signals, we replace r_e with $(r_e - r_a)$ to diminish the artificial effect along the edge of ROE, where r_a is the radius of the region of artifacts (ROA) centered at the origin. Let us assume that projections s_θ are known, the ROE is the subset of the projection with a disc radius of r_e centered at polar coordinates (r, θ_0). We can rewrite the function as follows:

$$(\text{ROE} - \text{ROA}) : \{p : p \in [r\cos(\theta - \theta_0) - (r_e - r_a) : r\cos(\theta - \theta_0) + (r_e - r_a)]\}. \quad (12.14)$$

The constant extrapolation satisfies

$$(s_\theta)_{\text{local}}(p) = \begin{cases} s_\theta(p) & \text{if } p \in \text{(ROE-ROA)}, \\ s_\theta(r\cos(\theta - \theta_0) + (r_e - r_a)) & \text{if } p \in [r\cos(\theta - \theta_0) + (r_e - r_a) : +\infty], \\ s_\theta(r\cos(\theta - \theta_0) - (r_e - r_a)) & \text{if } p \in [-\infty : r\cos(\theta - \theta_0) - (r_e - r_a)]. \end{cases} \quad (12.15)$$

Fig. 12.3(d) shows the extrapolated projection at the 25^{th} projection angle after the application of a scaling ramp filters and a ramp filter. The extrapolated projection removes spikes at the edge of the ROE. Fig. 12.3(e) and (f) shows the resultant sinograms after performing a scaling ramp filter and a ramp filter, respectively. The intensity variation in the two different filtered versions of sinograms is clearly seen.

It should be noted that there are phase shifts observed in Fig. 12.2(c) and (d) and Fig. 12.3(b) and (c), which are caused by the convolution of the ramp filter with the wavelet/scaling filter(s). However, since the wavelet and scaling filters are designed to satisfy the perfect reconstruction property, these shifts cancel out when the inverse 2D DWT is applied to yield the overall CT reconstruction.

However, according to the following local resultant reconstruction examples, it is observed that an approximate image has better resultant reconstruction than using inversed wavelet reconstructed subimages, in the current algorithm related to wavelet based reconstruction. Consequently we only consider the approximate image reconstruction to evaluate the algorithm validation. As we mainly deliver local information for the back projection, the wavelet and scaling ramp filters show essentially same support. But we have to recognise that having a scaling-only reconstruction means we theoretically lose some accuracy in locating any edges.

The local CT algorithm is also used for the reconstruction of the off-center region. The extrapolation algorithm is suitable for the reconstruction of the off-center region of interest. In this example, one portion of the sinogram with $1/3$ of all the projections is truncated for recovery of an image within the ROI and the extrapolated filtered projections occur along only one edge of the ROE, where nonlocal data are set to zero.

Fig. 12.4(a) and (b) shows projection profiles at the 25^{th} projection angle in the off-center area, before and after extrapolated projection, with the application of a scaling ramp filter and a ramp filter, respectively. It can be seen that the spikes at the edge of the ROE have been removed. Fig. 12.4(c) and (d) show the resultant sinograms corresponding to wavelet ramp filters and ramp filters at the off-centered region of interest.

Case study ♯ 2: Plastic vial target

A second experiment is performed on a simple sample with a nested structure, a tube inserted in a vial. The measured data are Fourier transformed and the phase of the Fourier domain responses is used to reconstruct the sample. Eq. (12.1) is used to reconstruct a local image using a scaling ramp filter, with G_{2^j} being multiplied by a shape scaling factor λ. This shape factor is used to smooth wavelet ramp filtered cropped sinograms. The Eq. (12.1) can be rewritten as,

$$I(x,y) = \int_0^{\pi} \mathrm{d}\theta \cdot \left[\int_{-\infty}^{\infty} \hat{s}(\theta,\beta)|\beta|\hat{g}_{2^j}[\beta(\cos\theta\cdot\lambda),\beta(\sin\theta\cdot\lambda)]\exp[i2\pi\beta\xi]\mathrm{d}\beta \right].$$

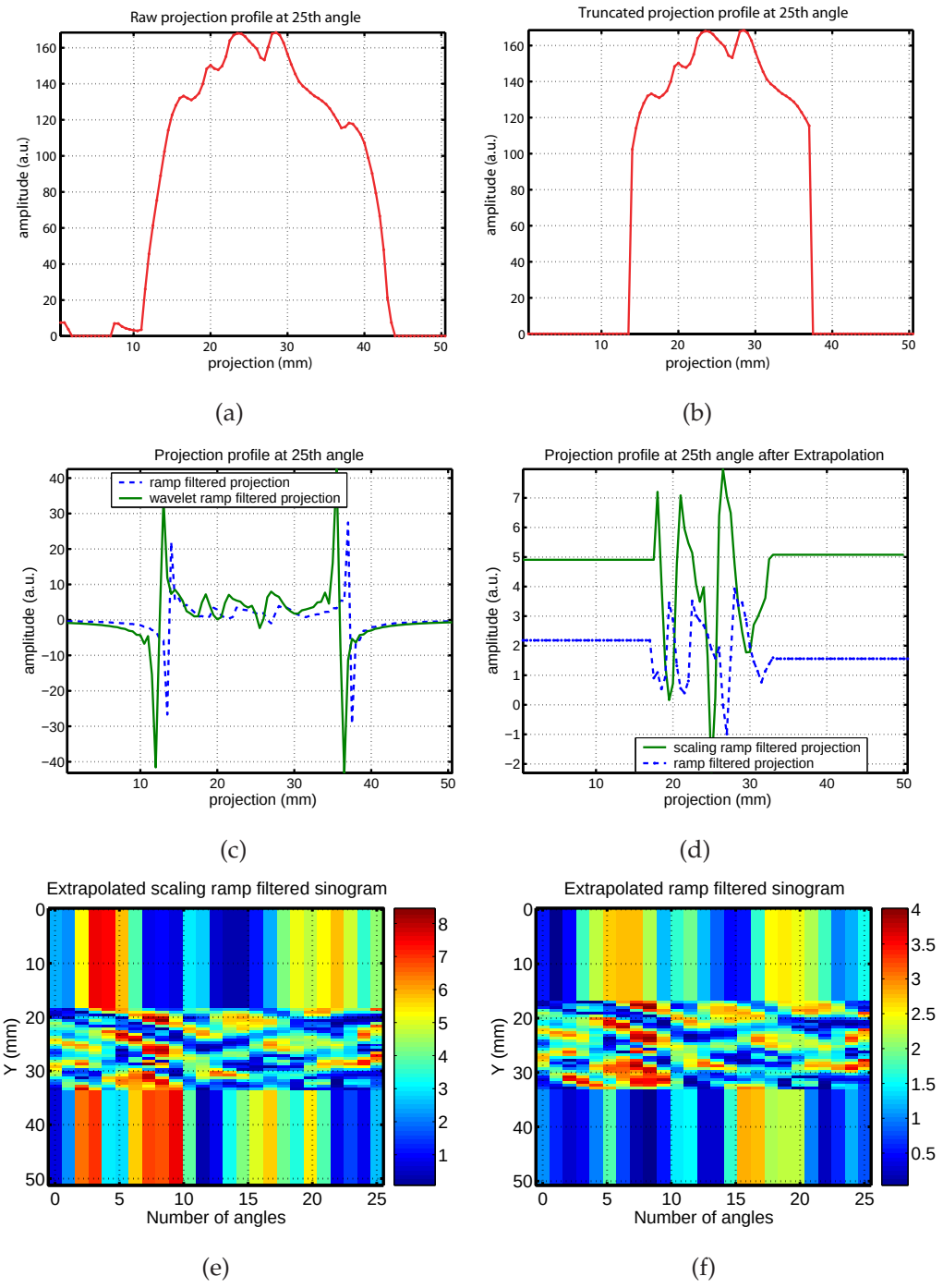


Figure 12.3. Wavelet based ramp filtered projections with extrapolation. (a) A sample projection with a polystyrene cylinder phantom with varying interhole distances. (b) Projection when nonlocal data are set to zero. (c) Projection filtered by a scaling ramp filter and a traditional ramp filter, respectively. (d) Projection extrapolation outside the ROI after filtered projections. (e) The resultant sinograms after scaling ramp filtered projections. (f) The resultant sinograms after ramp filtered projections.

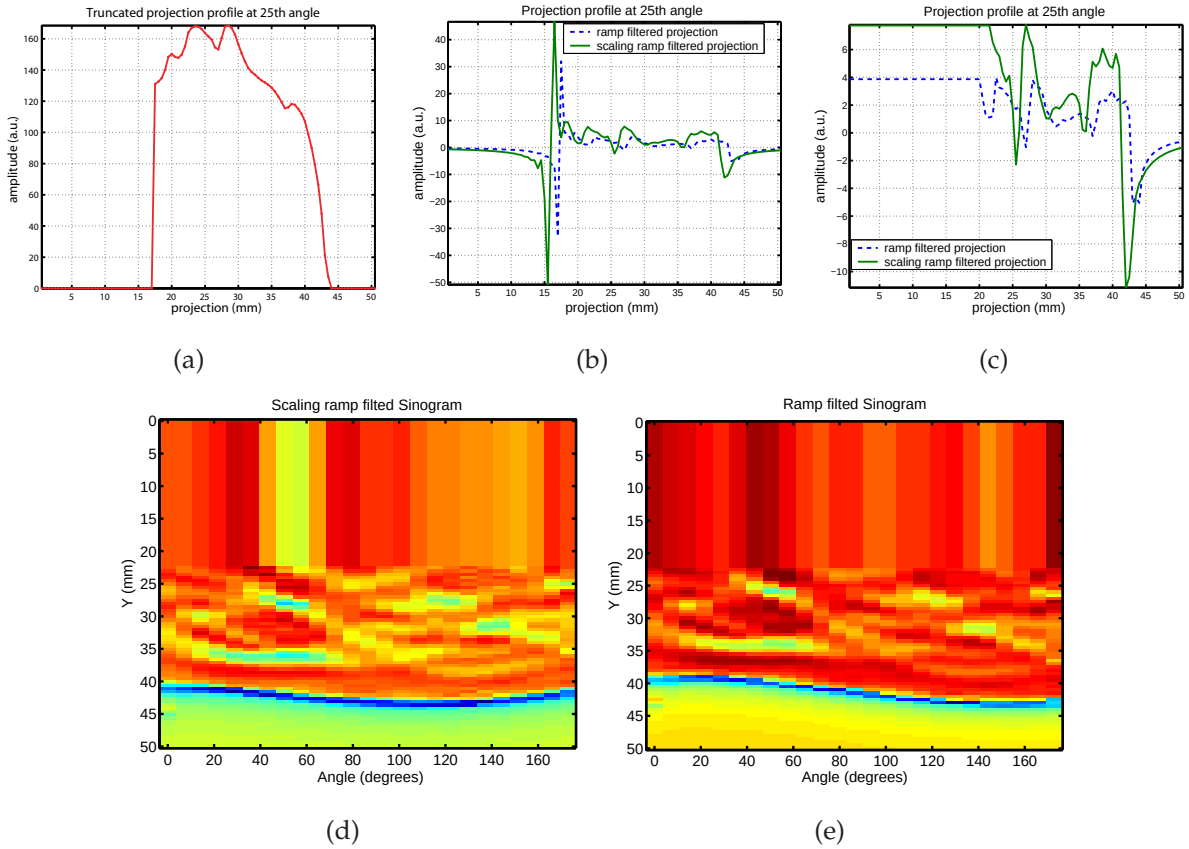


Figure 12.4. Projection in the off-center area. (a) Projection in the off-center area, when nonlocal data are set to zero. (b) Projection in the off-center area, filtered by a scaling ramp filter and a traditional ramp filter, respectively. (c) Projection of the off-center area extrapolated outside the ROI after filtered projections. (d) The resultant sinograms at the off-center area after scaling ramp filtered projections. (e) The resultant sinograms at the off-center area after ramp filtered projections.

Fig. 12.5(a) shows the wavelet ramp filtered projection at the first sampled frequency before a shape scaling factor is applied, where a large 'S' shape scaling ramp filtered projection is observed. Fig. 12.5(b) shows and almost flat border along the projection after applying a shape scaling factor of 1/3.

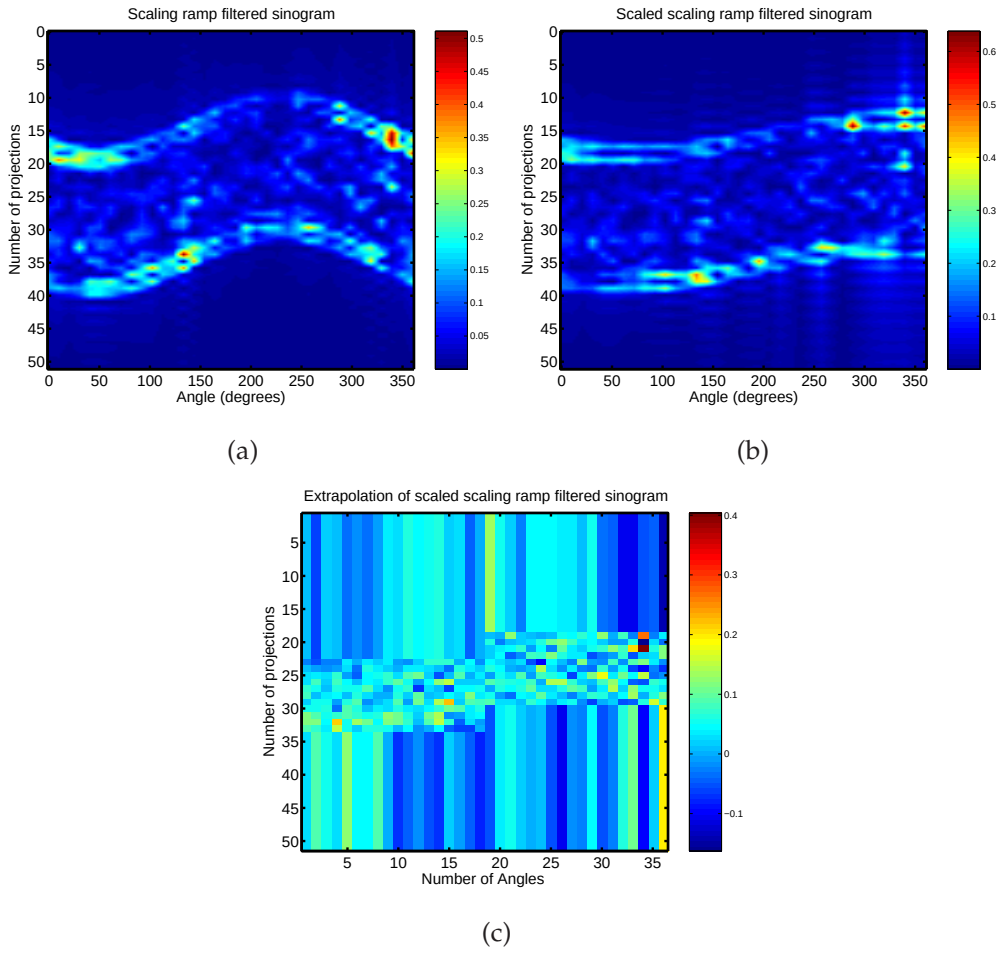


Figure 12.5. Wavelet based ramp filtered projections after applying a shape scaling factor of 1/3. (a) Scaling ramp filtered projection before a shape scaling factor is applied. (b) Scaling wavelet ramp filtered projection after a shape scaling factor of 1/3 is applied. (c) Illustration of the resultant sinogram via extrapolation of scaling wavelet ramp filtered projection.

In this example, illustrated in Fig. 12.5(b), the boundaries of the sinogram are two slowly ascend lines from left to right. Let us assume that the region of artifacts consists of two parts: ROA_1 and ROA_2 with the radii of r_{a1} and r_{a2}, and with the projection angles of $\theta_1 \in [0 : \theta_{a1}]$ and $\theta_2 \in [0 : \theta_{a2}]$, respectively, in an image. To overcome the problem of edge discontinuities, truncated regions in the sinogram are extrapolated with a constant value, in the manner described by Eq. (12.16). Fig. 12.5(c) shows the resultant sinogram via extrapolation of the scaling wavelet ramp filtered projection, with the same number of the projections at each projection angles being kept for convenience in calculation of reconstructed image:

$$
s_{\theta\mathrm{local}}(p) = \begin{cases}
s_\theta(p), \ \text{if } p \in (\text{ROE-ROA}_1), \text{and } \theta \in [0 : \theta_{a1}] \\
s_\theta(p), \ \text{if } p \in (\text{ROE-ROA}_2), \text{and } \theta \in (\theta_{a1} : \theta_{a2}] \\
s_\theta(r\cos(\theta-\theta_0) + (r_e - r_{a1})), \ \text{if } p \in [r\cos(\theta-\theta_0) + (r_e - r_{a1}) : +\infty], \\
\qquad\qquad \text{and } \theta \in [0 : \theta_{a1}] \\
s_\theta(r\cos(\theta-\theta_0) + (r_e - r_{a2})), \ \text{if } p \in [r\cos(\theta-\theta_0) + (r_e - r_{a2}) : +\infty], \\
\qquad\qquad \text{and } \theta \in (\theta_{a1} : \theta_{a2}] \\
s_\theta(r\cos(\theta-\theta_0) - (r_e - r_{a1})), \ \text{if } p \in [-\infty : r\cos(\theta-\theta_0) - (r_e - r_{a1})], \\
\qquad\qquad \theta \in [0 : \theta_{a1}] \\
s_\theta(r\cos(\theta-\theta_0) - (r_e - r_{a2})), \ \text{if } p \in [-\infty : r\cos(\theta-\theta_0) - (r_e - r_{a2})], \\
\qquad\qquad \theta \in (\theta_{a1} : \theta_{a2}].
\end{cases}
\tag{12.16}
$$

12.4.2 Algorithm summary

The wavelet based reconstruction algorithm assumes an image support of radius R, and the radius of the ROI is r_i. A radius $r_e = r_i + r_a$ is exposed, where r_a is the extra margin with related to radius of ROA, which is produced by applying wavelet filters on the project data. The algorithm is summarized as follows:

1. The original projections are calculated from time or frequency parameters from terahertz measurements.

2. The region of exposure is truncated for the reconstruction of an image in the region of interest.

3. The region of exposure of each projection is filtered by modified wavelet filters at all projection angles. This step is to recover an image related to wavelet detailed coefficients.

4. The region of exposure of each projection is filtered by modified scaling filter at all projection angles, which will lead to the recovery of the approximation sub-image. A shape factor λ is used to adjust the shape of sinogram in order to smooth the sinogram images.

5. The projections from Step 4 are extrapolated with tuning parameters to limit artifacts at the boundaries of the projections.

6. Filtered projections obtained in Step 3 and Step 4 are back projected to every other point to obtain the approximate and detail at the higher resolution. The remaining points are set to zero.

7. The image is reconstructed from the wavelet and scaling coefficients via a conventional inverse 2D DWT.

12.5 Reconstruction results

12.5.1 Case study ♯ 1: Polystyrene target

A 83×83 pixel image of the polystyrene target is recovered from the wavelet and scaling coefficients using global data, shown in Fig. 12.6(a), with interpolation in the inverse wavelet transform for clarity and comparison. Each measured terahertz pulse is a function of time with 401 samples at uniform time intervals of 0.067 ps. The time domain signals from the sample and reference pulses are truncated at the second sampled time step and applied for reconstruction. Wavelet and scaling coefficients after back projection are shown in Fig. 12.6(b), where the BioSpline2.2 biorthogonal basis is used. The quality of the reconstructed image is, as expected, almost indistinguishable from the reconstruction using traditional filtered back projection (FBP), shown in Fig. 12.6(c), with interpolation for an equal sized image as for wavelet based reconstruction. The differences between the wavelet based reconstruction and traditional filtered back projection are evaluated using the reconstructed profiles at the 40^{th} horizontal row of pixels and 40^{th} vertical column of pixels, illustrated in Fig. 12.6(d) and

(e), where it is not difficult to see minor variations in detected hole positions using wavelet version of reconstruction (dash line) compared to traditional FBP algorithm (dash dot line).

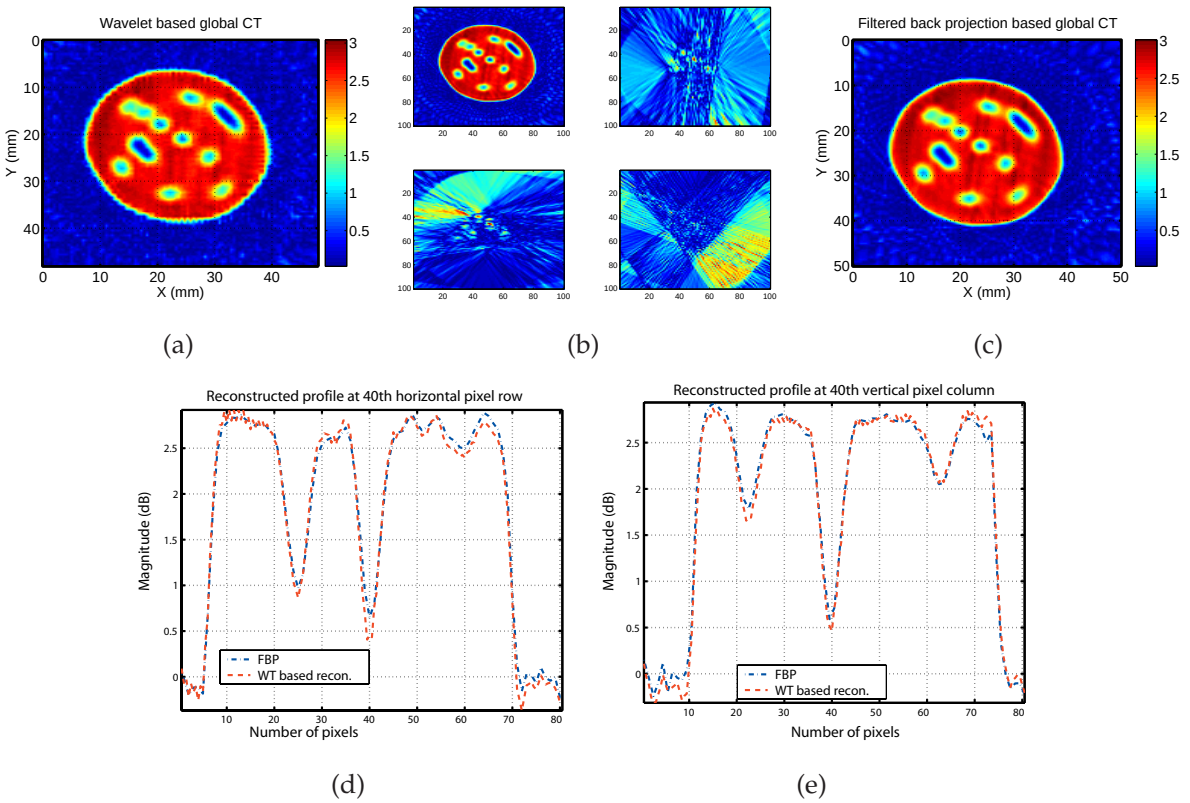


Figure 12.6. Wavelet and scaling ramp filtering of sinograms towards the polystyrene target. (a) A 83×83 pixel image of the polystyrene target is recovered from the wavelet and scaling coefficients using global data with interpolation for smoothness. This resultant reconstruction has the similar reconstruction quality to traditional back projection algorithm. (b) Wavelet and scaling coefficients after back projection. (c) A 83×83 pixel image of the polystyrene is reconstructed using traditional filtered back projection with interpolation for smoothness. (d) Reconstructed profiles at the 40^{th} horizontal pixel row. (e) Reconstructed profiles at the 40^{th} vertical pixels column.

Fig. 12.7 shows reconstructed images centered at a radius of 16 pixels using the local reconstruction method outlined in subsection 12.4.2 and the traditional FBP algorithm. Each reconstruction is evaluated on a 100×100 pixel image. Fig. 12.7(a) shows the truncated projections with 46% of full data. Fig. 12.7(b) is the local reconstruction after extrapolation from wavelet and scaling filtered projection, with downsampling. Fig. 12.7(c) shows four subimages reconstructed from wavelet and scaling coefficients

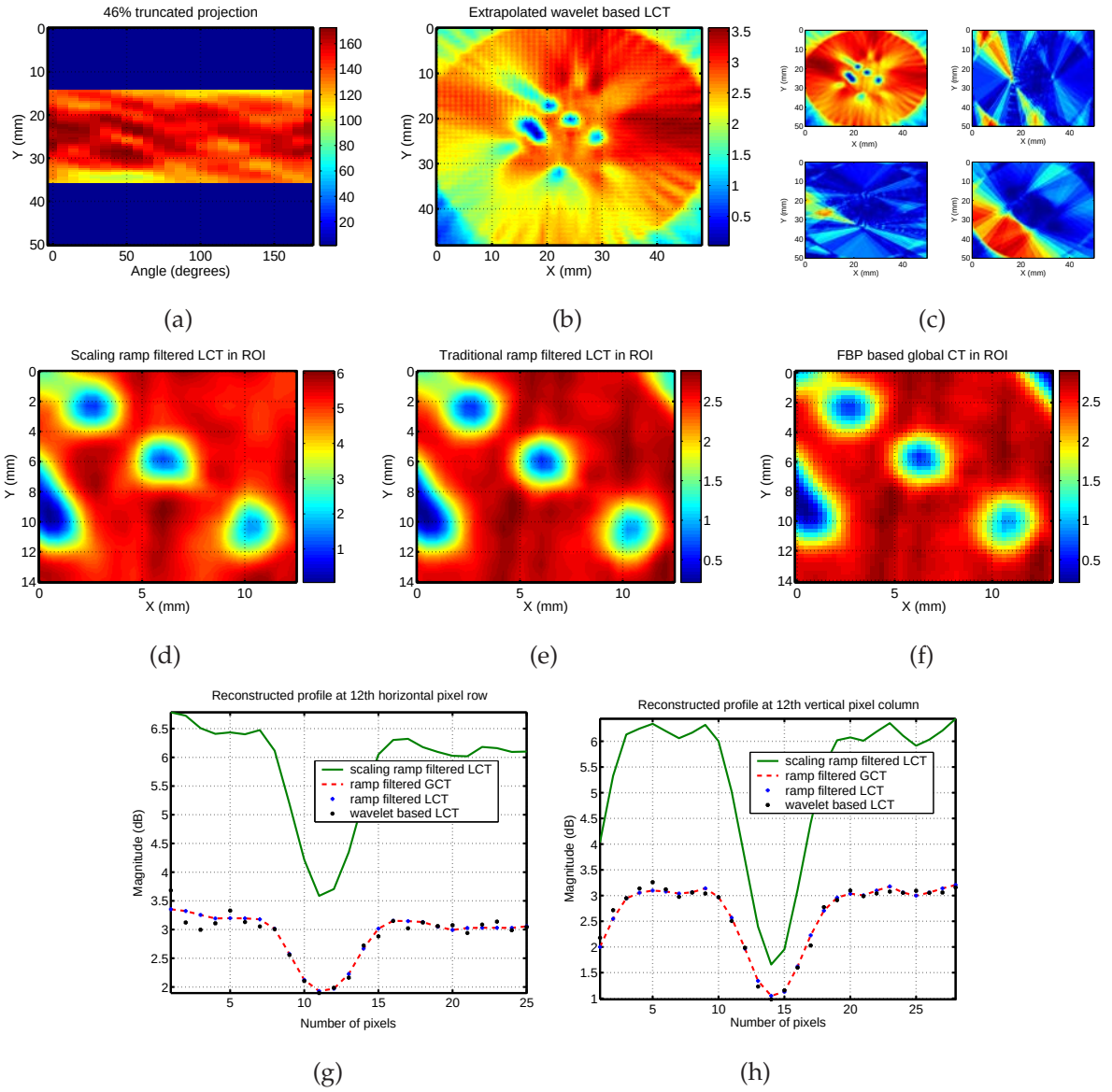


Figure 12.7. Reconstruction results of the polystyrene target with 46% of full data at centered area. (a) Illustration of the truncated projections with 46% of full data (b) Reconstructed image localised to a region of interest from the inverse wavelet transform. (c) Centered approximate and three detail reconstruction subimages along clockwise direction. (d) Cropped version of approximate subimage in (c). (e) and (f) A cropped reconstruction using traditional filtered back projection on local data and global data, respectively. (g) Reconstructed profiles at the 12^{th} horizontal pixel row. (h) Reconstructed profiles at the 12^{th} vertical pixel column.

after constant extrapolation and BP. The corresponding truncated version of the ROI is illustrated in Fig. 12.7(d), with a relative reconstruction error of 26% using our algorithm. A local reconstruction of the ROI using FBP is achieved in Fig. 12.7(e) for comparison, which results in a bigger relative error, proportional to the wavelet based error calculation. Fig. 12.7(f) is a truncated version of FBP based global CT. Fig. 12.7(g) and (h) shows the reconstruction profiles at the 12^{th} horizontal row and vertical column of pixels corresponding to each reconstruction. As illustrated in Fig. 12.7(g) and (h), the profile taken from the image reconstruction specific to wavelet scaling subimage shows obvious contrast in reconstructed intensity for different media relative to the remaining profiles from their reconstructed images.

Fig. 12.8(a)-(f) shows reconstructed images at an off-center area with a radius of 61 pixels using the current local reconstruction method and the traditional FBP algorithm. Each of the subfigures illustrates, for comparison, 33% of full projection data; local reconstruction from extrapolated wavelet and scaling filtered projection after decomposition; the reconstruction of extrapolated approximate and detail coefficients after BP, with a reconstruction error at the off-center ROI of 24%; traditional local and global reconstruction in ROI; the reconstruction profiles at the 28^{th} horizontal row of pixels and the 12^{th} vertical column of pixels are illustrated in Fig. 12.8(g) and (h), both of which correspond to the reconstructions from approximate wavelet coefficients, FBP based local and global recovery in the ROI. The reconstruction from wavelet approximate coefficients shows strong contrast in intensity for different media and FBP based local reconstruction shows a little higher intensity than FBP based global reconstruction.

According to the analysis mentioned above and the illustrations in Fig. 12.7(g) and (h) and Fig. 12.8(g) and (h), it is reasonable to apply a wavelet approximate subimage for the analysis of wavelet based LCT. On the one hand, because the filtered wavelets have essentially the same support, we need only local information to perform the back projection. On the other hand, it is a fact that a sharply, peaked scaling function may be very good at isolating small scale image features, such as high-frequency system noise, faint scattering and absorption around bright image features. It has been observed in these reconstructed images via wavelets.

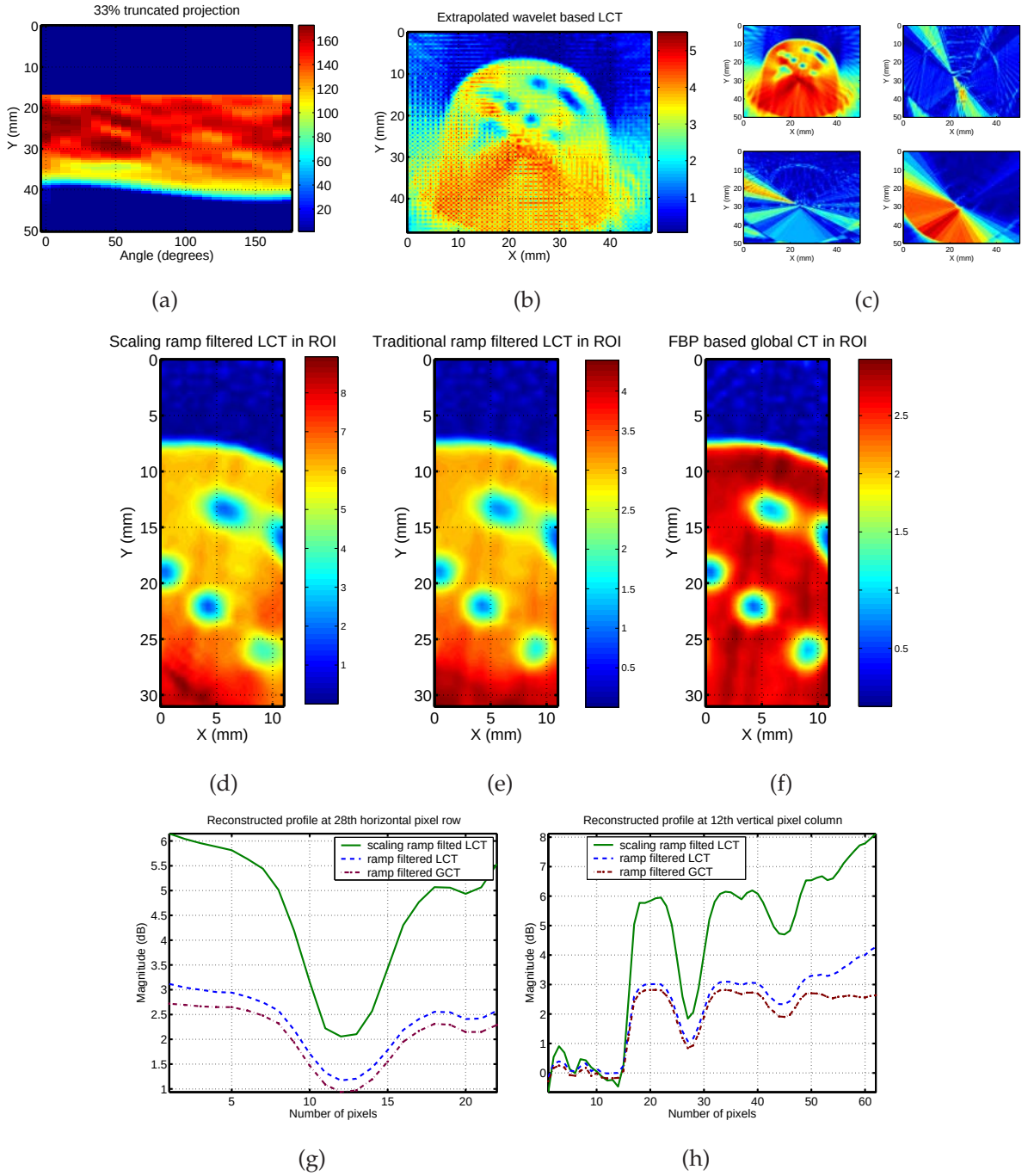


Figure 12.8. Reconstruction results of the polystyrene target with 33% of full data at off-centered area. (a) Illustration of the truncated projections with 33% of full data (b) A locally reconstructed image from the inverse wavelet transform. (c) Off-centered approximate and three detail reconstructed subimages along clockwise direction. (d) Cropped version of approximate subimage in (c). (e) A cropped of FBP based local reconstruction. (f) A truncated version of FBP based global reconstruction. (g) Reconstructed profiles at the 28^{th} horizontal row of pixels. (h) Reconstructed profiles at the 12^{th} vertical column of pixels.

12.5.2 Case study ♯ 2: Plastic vial target

The nested structure of a tube inside a PET vial is imaged on a 100×100 grid. Its reconstruction from the wavelet and scaling coefficients using global data is shown in Fig. 12.9(a). The ten images span the sampled frequency scope from ten lowest frequencies, from 0.0213 THz to 0.213 THz. Again, the BioSpline2.2 biorthogonal basis is used. The quality of the reconstructed image is similar to using traditional filtered back projection (FBP), shown in Fig. 12.9(b), with a little increased recovered image intensity in the reconstructed subimages and a little discontinuity in the third reconstructed subimage compared to the traditional FBP algorithm.

Fig. 12.9(c) and (d) shows reconstructed images after extrapolation, evaluated on a 100×100 grid, at a center area with a disk radius of 6 pixels using the current local reconstruction method and the traditional FBP algorithm. They are enlarged for clarity. Each of the reconstructed subimages is illustrated, from 0.0213 THz to 0.213 THz, relatively, with 59% of full projection data. The 59% of full projections is shown in Fig. 12.9(e) at the 6^{th} sampled frequency. The local reconstruction in the ROI from extrapolated wavelet and scaling filtered projection is shown in Fig. 12.9(c). Fig. 12.9(d) is the corresponding local reconstructions using FBP algorithm. The noise is reduced in wavelet based reconstructed images at the first two frequencies of 0.0213 THz and 0.0426 THz. It is valuable in the exploration of biomedical images using terahertz data. The wavelet approximate and detailed coefficients after BP at the 7^{th} sampled frequency is illustrated in Fig. 12.9(f), with a relative error of 29% from the approximate reconstruction.

12.6 Future work

Since the current work involves only the one level of 2D DWT, it will be interesting to explore the reconstruction algorithm with more levels of decomposition. Moreover, a research area of much current interest is the development of statistical based local tomography algorithm and techniques (Kolehmainen *et al.* 2003, Hanson and Wecksung 1983). It aims towards the actual localised reconstruction with relation to the terahertz measurement. The wavelet technique is useful for local reconstruction, and the relative wavelet transform coefficients can be thresholded to reduce the computation complexity (Meyer-Base 2003). In addition, the current resultant experiment

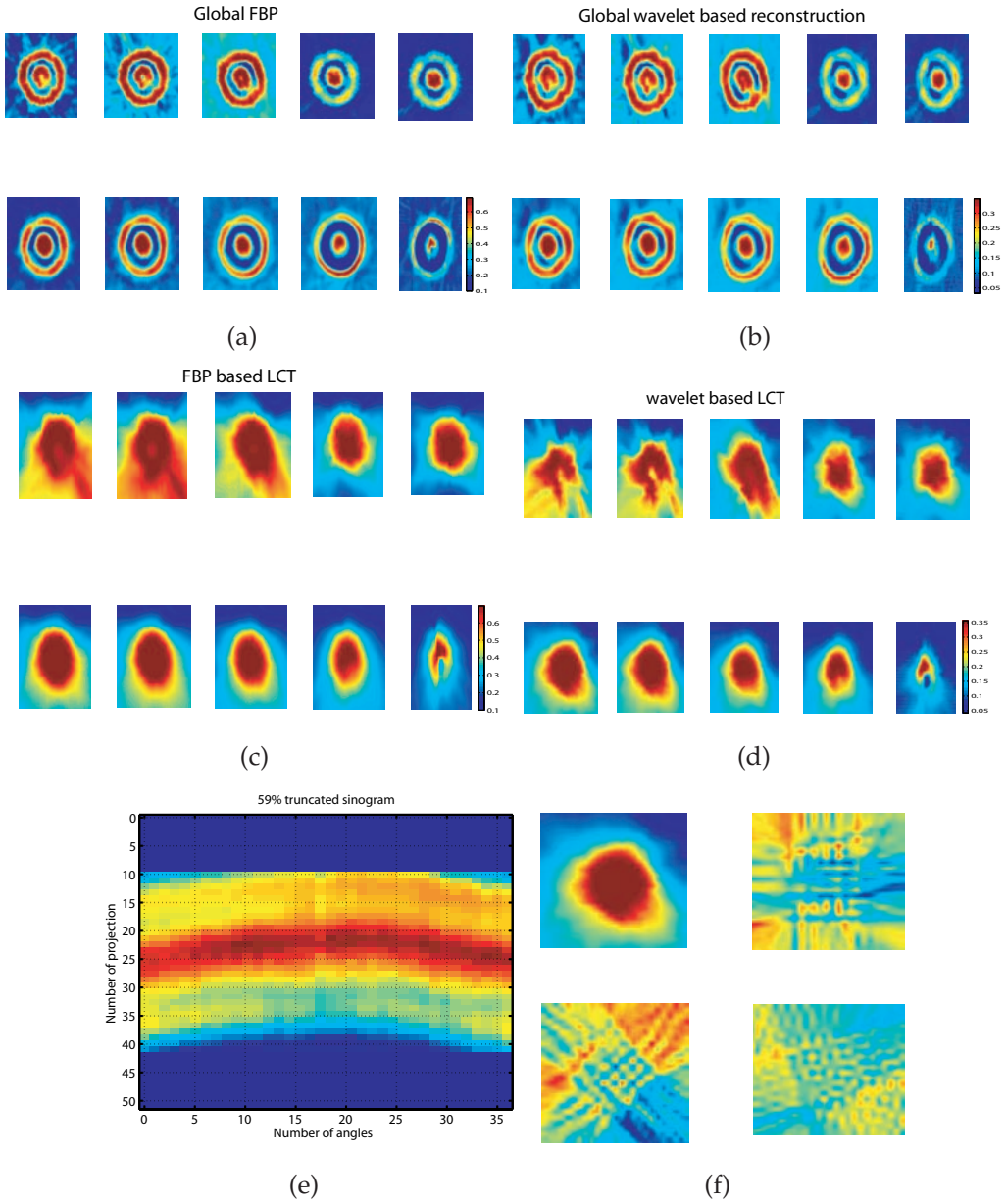


Figure 12.9. Reconstruction results of a vial slice from the nested structure. (a) Illustration of a 100×100 pixel global image of the tube inside a vial, with frequency range from 0.0213 THz to 0.213 THz. It is recovered from the wavelet and scaling coefficients, after decomposition. (b) FBP based reconstruction from global measurement with size of 100×100 pixels and the same frequency range of (a). (c) Corresponding reconstruction FBP algorithm using local projection data. (d) A reconstructed image of the tube from the inverse wavelet transform after decompostion. (e) Illustration of the truncated projections with 59% of full data. (f) Approximate and detail reconstruction coefficients after BP using local projection data.

relies on the fact that the Hilbert transform (part of the inverse RT) does not really change the compact support of scaling and wavelet functions.

12.7 Chapter summary

We have developed an algorithm to reconstruct the wavelet and scaling coefficients of a function from its Radon transform of terahertz signals. Based on the observation that for some wavelet bases, with sufficient zero moments, the scaling and wavelet functions have essentially the same support after ramp filtering. An upper bound for the local reconstruction error is obtained in terms of the amount of nonlocal data, which is used in the reconstruction scheme. Two targets are recovered from terahertz measurements, which demonstrates the current local reconstruction methods using a wavelet based transform scheme.

The next Chapter extends the application of 2D wavelet based local reconstruction to a 3D space, with the use of a narrow THz source, a THz quantum cascade laser emitting at 2.9 THz. The continuous wave (CW) THz QCL is experimentally illustrated in the next Chapter, further illustrating the application of wavelet transforms for THz computed tomographic reconstruction. The resultant segments from the local reconstructed images are compared with the ground truth to explore the ability of a QCL to image target object, polystyrene, with complex contours. It is an evidence that local computed tomography via wavelets is suitable for the image reconstruction in terahertz frequency range, which results in lower misclassification after segmentation, than with traditional FBP algorithms.

Chapter 13

Local CT Using a THz QCL

WE study the local reconstruction of a region-of-interest (ROI) from a 3D terahertz data set, obtained via a quantum cascade laser (QCL). Difficulties with the limited projection angles and image noise make the development of accurate local reconstruction algorithms particularly challenging.

Segmentation algorithms are applied on the reconstructed images with low contrast. The resultant segments from the local reconstructed images are compared with the ground truth to explore the ability of a QCL to image a polystyrene target object, polystyrene, with complex contours. In this Chapter, we use a polystyrene clown's head with a hole inside as a target.

It is found that 3D local reconstructions of the target (hole) using a QCL take on a number of different shapes since the various contours of the target physically distort the measured optical parameters of the object. Local computed tomography via wavelets is found to be more suitable for the image reconstruction in terahertz frequency range in the case of lower image quality. The wavelet based scheme was able to produce a lower misclassification rate than the FBP-based algorithms.

13.1 Introduction

The potential of high power THz QCLs for biomedical imaging and security detection is significant (Kim *et al.* 2004). Terahertz imaging either depends on terahertz pulsed imaging (TPI) or continuous wave (CW) terahertz imaging. Time-domain TPI has the advantage of providing a broad frequency spectrum, and resolving the arrival of the photons as a function of time at different locations around the specimen boundary. In contrast with THz pulsed imaging, QCLs afford high average power sources of continuous wave (CW) radiation to achieve deeper penetration of samples, which has been identified as one of the principal challenges for terahertz tomography (Zhang *et al.* 2004). The CW terahertz imaging methods have the difficulty of resolving target material absorption from scattering (Ntziachristos *et al.* 2005). We have to recognize that there are still challenges for QCLs to reach higher operating temperatures and high signal-to-noise ratio (SNR), nevertheless, several tens of milliwatt peak terahertz powers (Barbieri *et al.* 2004) and a few tens dB of SNR can already be very useful for biomedical imaging and other applications. Meanwhile, QCLs retain coherent detection properties. The Thesis aims to investigate the scattering of terahertz CW radiation with a QCL via applying proposed reconstruction algorithms to the sample with complex contours.

Quantum cascade lasers (QCLs) are semiconductor injection lasers, based on quantum semiconductor structures that are grown by molecular beam epitaxy and designed by band structure engineering (Faist *et al.* 1994). Quantum cascade lasers, in principle, are realised via intersubband transitions in a multiple-quantum-well (MQW) heterostructure (Gmachl *et al.* 2001, Tonouchi 2007). In contrast to the conventional diode laser, QCLs only involve electron transitions that occur between the conduction bands (intersubband) instead of from the conduction band into the valence band (Köhler *et al.* 2001). As a result, the emission wavelength can be controlled by the thickness of the MQW, and the intrinsic high-power capabilities of the lasers make possible the cascading process, in which each electron generates several tens of photons in superlattice structures (Tredicucci *et al.* 2005). In addition, intersubband transitions are characterized through ultrafast carrier dynamics and band-structure engineering is available to successfully control the electron flow and thus increases population inversion, which ultimately controls the laser threshold (Gmachl *et al.* 2001, Ozyuzer *et al.* 2007).

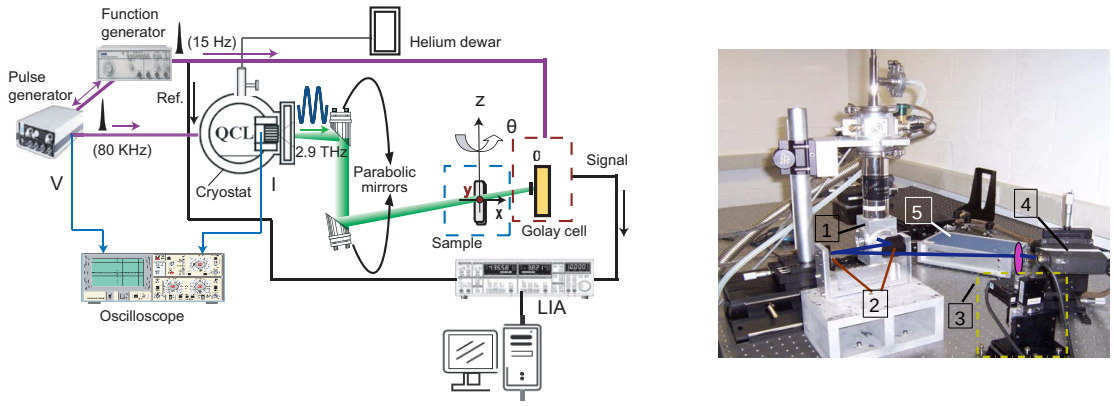
Terahertz QCLs were first reported only recently (Köhler *et al.* 2001). This is because an insuperable barrier exists in the phonon reststrahlen band (in the thermal infrared

region, refers to strong polar phonon absorption of energy), which causes difficulty in the generation of lower terahertz frequencies. Moreover, the development of a suitable waveguide (Sirtori *et al.* 1998a) is necessary to confine the long wavelength T-rays to an epilayer with low absorption losses by applying molecular beam epitaxy technology (Tredicucci *et al.* 2005). New design concepts have led to continuous-wave operation and high output powers (Rochat *et al.* 2002). Recently, terahertz quantum cascade lasers have made rapid progress. Examples range from chirped superlattices (Köhler *et al.* 2002a) to the surface plasmon concept (Köhler *et al.* 2002b) employed for large optical confinement without imposing high propagation losses. Recently, the technique of bound-to-continuum transitions and extraction of carriers via resonant phonon scattering (Williams *et al.* 2003) has been used more frequently. In addition, progress concerning terahertz single-mode operation and frequency extension has also been reported (Tredicucci *et al.* 2005).

It should be emphasized that unlike X-ray imaging, which realises parallel scanning using multiple point detection, point-to-point detection is the fundamental scheme of terahertz scanning, though it is non-ionising compared to X-ray imaging. The main goal of this Chapter is to present a wavelet based reconstruction algorithm and a standard reconstruction techniques for terahertz computed tomography using a QCL imaging system and 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. Recall from Chapter 1 (pp. 6-7) that the rationale for using wavelets in the context of a CW system, is that we are performing *local* tomographic reconstruction, here, where we take advantage of wavelet localization to enable zooming into locally reconstructed features. We illustrate the filtered sinograms and reconstructed images in the region of interests at four different measurement heights, using a clown's head with a hole inside as the target. The error is calculated via comparing segmentation of image reconstruction using FBP algorithms and wavelet based algorithms with the segments from the ground truth, respectively. The resultant segments are evaluated by calculating the slope of each two centroids at two different slices. The results show that wavelet based LCT has relative lower error in segmentation than FBP based LCT, and even better than global computed tomography (GCT) using FBP algorithms, though there is a slightly changed linear structure of the target using wavelet based local tomography compared to FBP based global CT and local CT. Our results also show varying degrees of optical distortion due to light scattering resulting from these various contours of the

different slices within the specimen. Meanwhile, the current reconstruction is proof-of-concept that it is necessary to develop a terahertz pulsed quantum cascade laser (McManus *et al.* 2005)—high resolution laser spectroscopic techniques—for image recovery.

13.2 A T-ray quantum cascade laser



(a) (b)

Figure 13.1. Terahertz quantum cascade laser setup. (a) Experimental apparatus for a terahertz QCL imaging system that is used to realise terahertz CT imaging. In practice, the sample is placed on a rotational stage for multiple view angles, which is mounted on a x-y-z linear stage to perform 3D scanning. (b) The photograph of a part of the terahertz QCL imaging system. The numbers in the text-boxes from 1 to 5 indicate, respectively, a QCL that is mounted on the cold finger of a continuous-flow helium-cooled cryostat, a pair of parabolic mirrors, the rotational stage mounted on an x-z translational stage, a Golay cell detector, and a detector controller.

As a visiting scholar, the author accessed the QCL setup described herein at the University of Cambridge. A schematic diagram of the current quantum cascade laser imaging apparatus is illustrated in Fig. 13.1(a). The pulse generator supplies pulses at a frequency of 80 kHz to the QCL. The output signal from the pulse generator is usually gated with a 15 Hz, 50% duty cycle slow modulation by an electronic chopper (a function generator of pulses) to match the detector (the Golay cell) response time, and to afford a reference frequency to the lock-in amplifier (LIA). The LIA is used to digitise signals and to significantly improve the signal-to-noise ratio by setting a time constant,

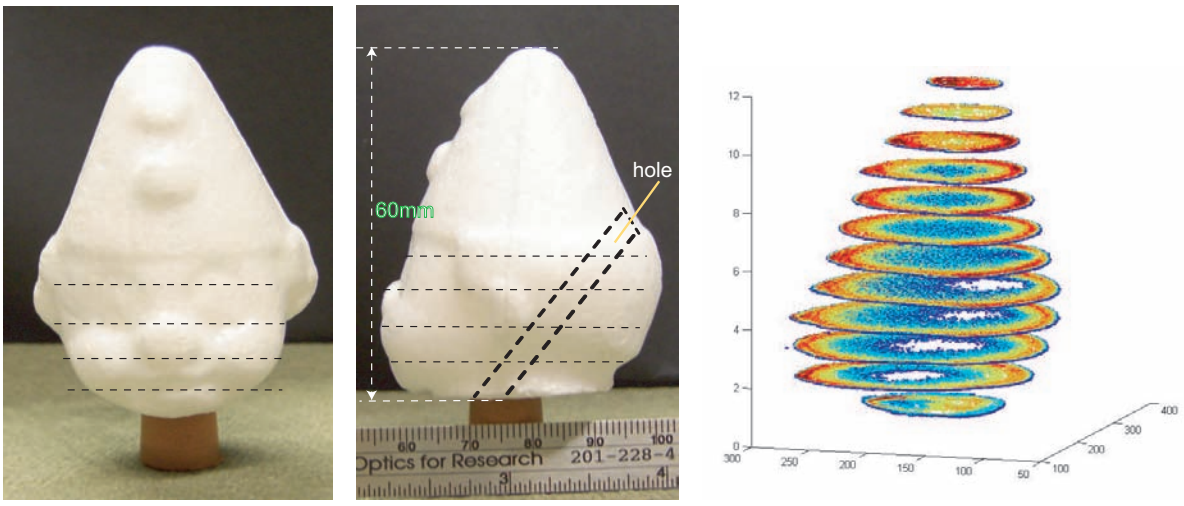
over which the input signal is integrated for each data point. The optimal time constant is set to 50 ms and a threshold current density is set as 112 A/cm^2.

The current terahertz QCL is a GaAs-AlGaAs bound-to-continuum superlattice design, emitting at 2.9 THz (103 μm), and grown by molecular beam epitaxy (Barbieri *et al.* 2004). This device typically operates up to 95 K in pulse mode and delivering 70 mW per facet peak power. The QCL is mounted on the cold finger of a continuous-flow helium-cooled cryostat maintaining a heat-sink temperature of 4.2 K. The emission is collected with a 2″ f/1 off-axis parabolic mirror, then focused by a 2″ f/6.43 parabolic mirror onto the sample. The sample is mounted on a rotational stage, which is itself mounted on a translational stage. This current terahertz imaging setup achieves point to point detection. A linearly moving stage, labelled by x axis, enables parallel scanning with linear velocity of 20 mm/s and a rotating stage, labelled by θ, allows projections of an object to be taken at a number of projection angles with a rotary velocity of 4 degree/s. Another axis, labelled by y, is perpendicular to x axis, for a transform of two sets of coordinates and image reconstructions. To perform a full 3D imaging, another translational stage, labelled by z axis, allows the sample to move vertically in order to obtain images of cross sections at various heights. The transmitted beams are detected by a Golay cell. The power incident on the sample, including the effects of the transmission of the cryostat window, is typically $\sim$35 mW (peak). The signal-to-noise ratio in the absence of a sample in the beam is 20 dB (Nguyen *et al.* 2006). The photograph of a part of the terahertz QCL imaging system is shown in Fig. 13.1(b).

13.3 Implementation

13.3.1 Experimental Considerations

The THz local reconstruction is an extended projection according to the resultant experiment conducted by Nguyen *et al.* (2006). In this experiment, one set of terahertz QCL intensity data is considered for Local Computed Tomography (LCT): a nested structure of a polystyrene clown's head with a hole inside, see the target photo in Fig. 13.2(a) and (b). The target sample is imaged in 12 slices, from bottom to top at twelve different heights, 5 mm apart. The diameter of the hole is 10.1 ± 0.2 mm, measured directly from the target. For the local reconstruction, only the second to the fifth image slices are considered with the hole going through at a tilted angle of 43° to the vertical. The center of the hole at the first slice is also centered at the bottom cross-section. The first



(a) (b) (c)

Figure 13.2. Photographs and image reconstruction of a clown's head with a hole inside. (a) The photograph of a nested structure of a clown's head with a hole inside. (b) The photography of the side face of the clown's head. (c) The reconstructed slices are stacked to recover the nested structure of the target object with a hole inside (Nguyen *et al.* 2006). The lines in (a) and (b) indicate the current four sliced heights for LCT in the hole area of interest.

three target layers (from the second to the fourth image slice), labelled Slices 1 to 3, are imaged using 289 projections at each of 18 projection angles covering a 180° projection area to produce a 250 × 250 pixel image (edge truncated), while the fifth slice (Slice 4) is imaged using 268 projections at the same 18 projection angles. Fig. 13.2(c) shows the reconstructed slices, which are stacked to recover the nested structure of the target object, where the lines indicate that four sliced heights which are used for the exploration of the current local reconstruction of the hole geometry.

13.3.2 Error analysis

In order to test the performance of the local imaging system, only binary grey-level images are used via segmentation techniques, after back projection processes, including (1) a truncated version of global CT via FBP, (2) LCT via wavelet transforms, and (3) LCT via FBP, are carried out in the region of interest. The pixel count difference is calculated between the ground truth and each of the resultant segments at four different measurement heights. The error ratio is regarded as the difference divided by

the total number of pixels of the binary grey-level image from the target sample. For further validation of our segmentation scheme, slopes computed from the centroids of the segments at two different slices are used to compute the error of our reconstruction, by comparing with the true slope of the linear 3D (hole) structure. It is noticed that, as the material of the target phantom is expanded Styrofoam, the surface of the walls inside the hole is rough. But we consider that, since the roughness is uniform, compared to the variety of the complex contours, the effect caused by the roughness on reconstructed segments can be ignored.

Characteristics of the hole image

A typical local reconstruction, containing a single hole, is shown in Fig. 6(a). The intensity histogram of the relative image is depicted in Fig. 6(b).

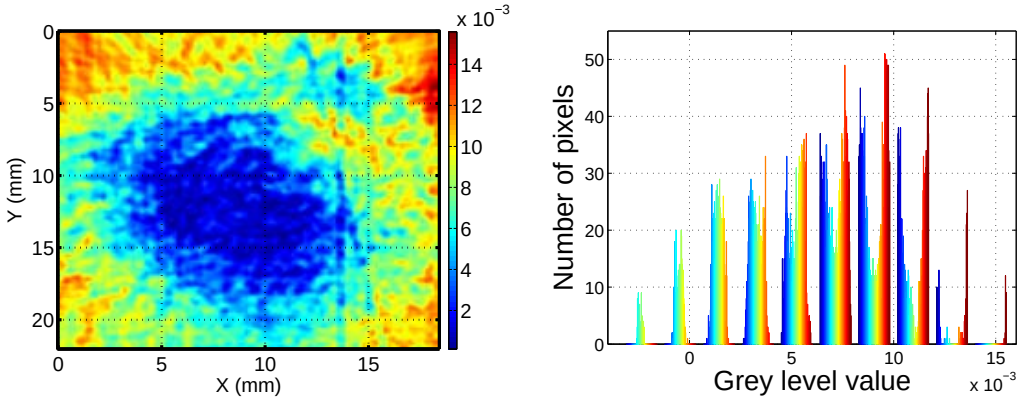


Figure 13.3. Illustration of a typical local reconstruction. (a) Illustration of a typical local reconstruction of a region of the target sample containing a single hole. (b) The intensity histogram of the relative image of (a) is depicted.

The typical local reconstruction concerns a single hole through the sample. There are several notable characteristics of this image, which are common to local reconstructions of the hole embedded in a 3D target at the different slices by different image reconstruction algorithms. Firstly, the contrast of intensities between the hole, the clown's head slice (object) and background is low. This is evident from the intensity histogram in Fig. 13.3(b). However, the average intensity of the object and background is often roughly uniform, separately, with a slightly darker region for the hole embedded on

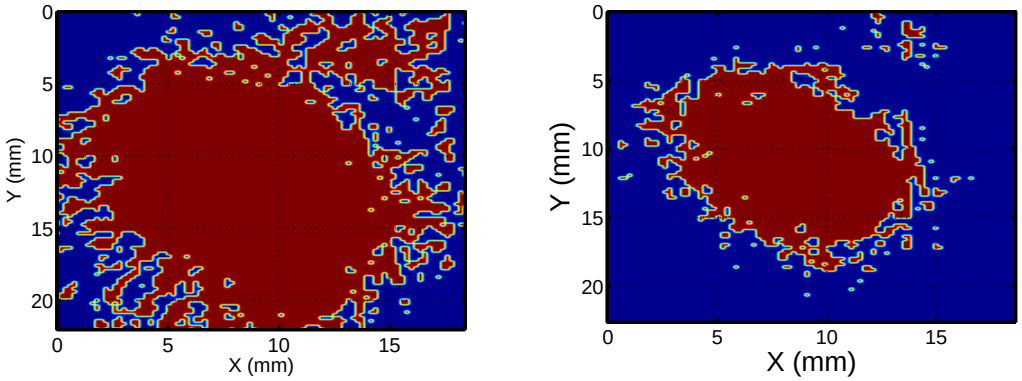


Figure 13.4. Segmented image via fuzzy c-means thresholding. Illustration of the result of applying two calculated thresholds via the fuzzy c-means thresholding method on the image of Fig. 13.3(a). (a) The image between the object region and background, achieved via using one of the two calculated thresholds. (b) The segmented image achieved via using the other threshold.

target across-section. Secondly, it is observed that the hole embedded in the target physically distort the measured optical parameters of the target medium that are used to reconstruct tomographic images correctly. The extent of shape distortion of a target media of interest on measured intensity depends on the slice contour and hole distance from the target boundary at each different measurement height. The varying exterior boundaries of the different slices result in varying degrees of optical distortion due to light scattering. In addition, a longer path length from the hole center position at each slice to the laser results in larger absorption. The 3D local reconstruction of the hole target takes on a number of different shapes in the different slices. As a result, the boundary of the hole is potentially difficult to define. Due to variations in intensity, both within the hole position and in the background, portions of the hole appear to blend directly into the background, without creating a distinct boundary.

Segmentation strategy

The aim of segmentation is to separate an image into different meaningful regions, by identifying the regions in an image with common properties and distinguishing dissimilar regions (Levine 1985). A popular segmentation scheme is image thresholding

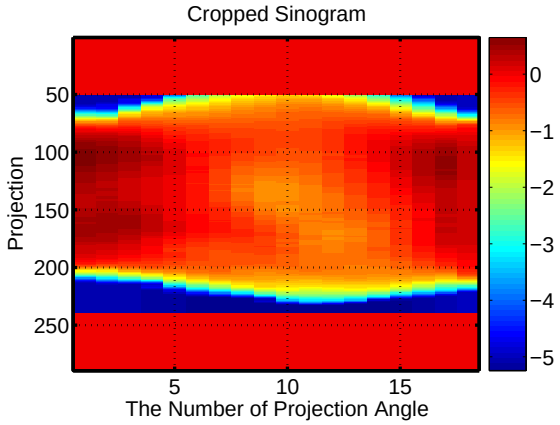


Figure 13.5. A cropped sinogram image. Illustration of the truncated projections with 79% of full data at the center area of interest.

(Haralick and Shapiro 1985, Gonzalez and Woods 2002), which can be regarded as a form of per-pixel classification. Grey-level intensity is used as a feature value to characterise each pixel of an image. The calculated threshold is compared with the feature value in order to map the pixel onto one of two groups, object and background regions.

In this work, two threshold levels are used to segment the entire image, as the recovered image, e.g. Fig. 13.3(a), shows a gradual change in intensity from the three regions: the hole, the clown's head cross-section, and the background. Since separation of different clusters is not well-defined, with dependence on the local grey value distribution, a fuzzy c-means technique is applied in this work, the selected set of feature vectors are strictly local in nature, that is, they depend on the 8-neighborhood of the pixel concerned. Compared to traditional Otsu's thresholding, the fuzzy c-means (FCM) method is effective for automatically computing two or more optimal segmentation thresholds, and for achieving the separation of underlying distributions. A fuzzy c-means threshold provides an iterative measurement to minimize the classification error allowing each pixel to be the member of all the possible classes with varying degrees of membership.

According to (Mukherjee *et al.* 1996, Chuang *et al.* 2006), the FCM algorithm assigns pixels to each category by using fuzzy memberships. Let $X = (x_1, x_2, ..., x_N)$ denote an image with N pixels to be partitioned into c clusters. The algorithm minimizes the cost function iteratively to find the best location for each of the clusters. The cost function

satisfies a least-squared error criterion

$$J = \sum_{j=1}^{N} \sum_{i=1}^{c} u_{i,j}^{m} ||x_j - v_i||^2 \tag{13.1}$$

where $u_{i,j}$ represents the membership of pixel x_j in the i^{th} cluster; v_i, whose number of components depends on the number of feature vectors (Mukherjee *et al.* 1996), is the cluster center of i^{th} class of fuzzy c-partitions, also called seed point for the particular class. The results of the calculated error or cost function are updated in every iteration according to $u_{i,j}$ and the distance $||x_j - v_i||$. This continues till the cluster centers become stable without significant difference between cluster centers in two consecutive iterations. The weighting exponent m is a constant with m equal to or bigger than 1, which controls the fuzziness of the resulting partition. The exponent of $u_{i,j}$ equal to 2 and maximum number of iterations equal to 100 are studied. The cost function is minimized when pixels close to the centroid of their clusters and are assigned high membership values, and *vice versa*. The membership function represents the probability that a pixel belongs to a specific cluster. In the FCM algorithm, the probability is determined by the distance between the pixel and each individual cluster center in the feature domain. The membership functions and cluster centers are updated according to the following:

$$u_{i,j} = \frac{1}{\sum_{k=1}^{c} (\frac{||x_j - v_i||}{||x_j - v_k||})^{2/(m-1)}} \tag{13.2}$$

and

$$v_i = \frac{\sum_{j=1}^{N} u_{i,j}^{m} x_j}{\sum_{j=1}^{N} u_{i,j}^{m}}. \tag{13.3}$$

The iterative procedures of FCM start from an arbitrarily assigned initial cluster center v_i, which leads to a solution that FCM converges to. For this experiment, we select a 3-class fuzzy c-means clustering to simplify the program process, therefore, there is no convergence involved. Fig. 13.4(a) and (b) are the resultant segments between the hole region and the region of object (including background), as well as the segments between object region and background, using two different calculated thresholds, separately. In practice, it is found that the threshold calculated via FCM results in loss of pixels in the hole region. We adjust the resultant threshold based on the ground truth to achieve the least error in the number of pixels between the resultant segment and the ground truth. In this work, since we only consider a local segment of the hole region, we focus on thresholding just for the hole, which is followed by post-processing

to clean up the segmented region. For convenience, in the following, we call the region of object adjoined to background as generalised background.

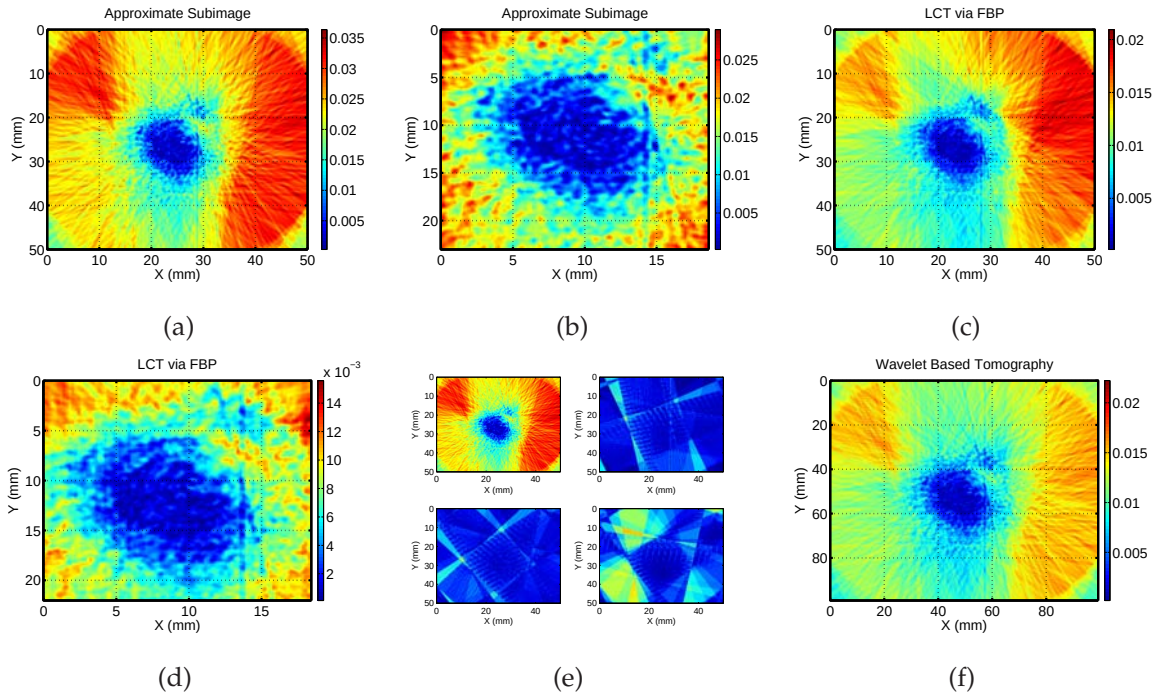


Figure 13.6. Resultant local reconstructions at Slice 1. Illustration of the resultant local reconstruction at Slice 1. (a) Reconstructed image is localised to a region of interest from the scaling function after constant extrapolation. (b) Cropped version of approximate subimage in (a). (c) Reconstructed image is localised to a region of interest via traditional FBP. (d) Cropped version of (c). (e) Centered approximate and three detail subimages for reconstruction along clockwise direction. (f) Reconstructed image is localised to a region of interest from the 2D inverse wavelet transform with interpolation.

Generally, thresholding an image can incur two kinds of noise: background pixels misclassified as a hole pixel, which produce small disjoint hole components in addition to the hole region, and hole pixels misclassified as the background, which produce gaps in the hole region of interest, see Fig. 13.4(b). Misclassification of either target sample of interest or background pixels near the hole boundary can also produce an extremely coarse boundary. In order to obtain the boundary representation of the hole region, region-based segmentation operations and morphological operations (Gonzalez and Woods 2002) are combined to repair the gaps in the hole region of interest and improve the smoothing of boundaries. This approach aims to apply *region growing* on

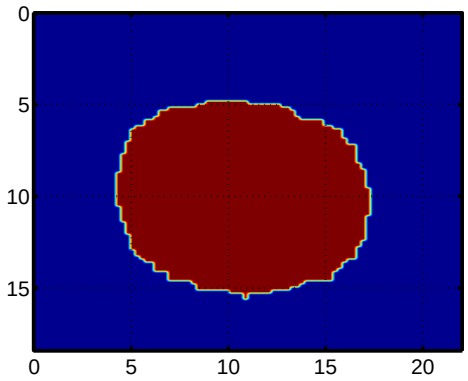


Figure 13.7. Plot of the measured ground truth. The ground truth is computed from the physical measurements of the target and using the planar projection of the tubular hole structure—ellipses with semimajor axes of 13.2 ± 0.6 mm and semiminor axes of 10 ± 0.8 mm along with the rotation angle of $8 \pm 2^{\circ}$.

a typical local reconstruction, which is represented as follows. Eight neighbor pixels are used as *seed pixels*. There are four starting points from four sets of seed pixels, which are applied for segmentation. Those starting points are positioned at the central point of the LCT in ROI or the point around the center. The four sets of seed pixels move from pixel to pixel along four quadrants of the coordinates, separately, starting from the four starting points. Predefined criteria are selected depending on the broken characters that are shown in each reconstruction. The maximum length of the breaks in the eight neighboring pixels is set to be two to four pixels. During the processing, we assign 1 to the pixel satisfying the selection criteria, otherwise we set the pixel to zero. This region growing is conducted repeatedly till a single connected hole region of interest appears without gaps.

In some cases, artifacts can be mistaken for the ROI (i.e. the hole). After region growing is used on regions obtained by thresholding the local reconstruction, a simple criterion is adopted to examine the length feature of the segmented hole region of interest, to ensure that the correct hole target region is selected. In this way, large artifacts can be eliminated, and small artifacts can be removed by region growing.

To calculate the error of segmentation based reconstruction, we compare each of the resultant segments with the ground truth. The error ratio is computed as the number of different pixels between the resultant segment and the ground truth divided by the number of pixels of the segment from the ground truth. The ground truth is computed

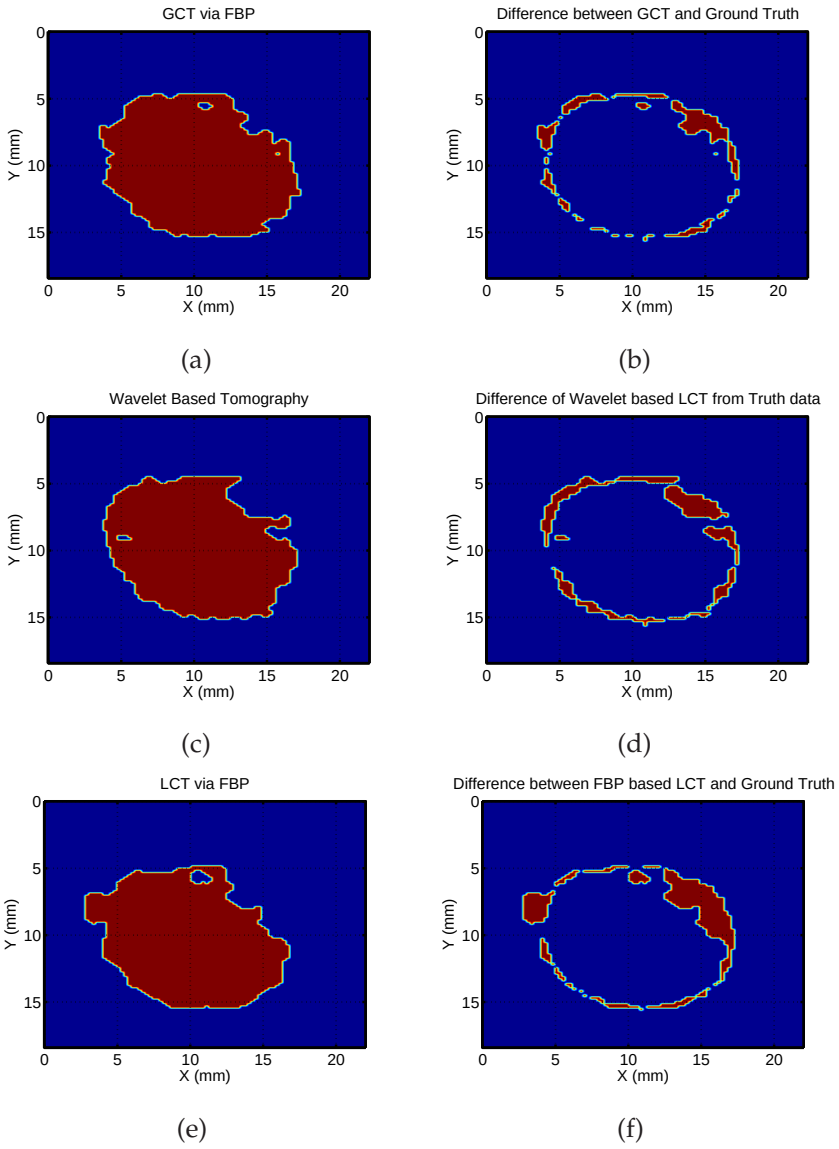


Figure 13.8. Resultant segments and differences at Slice 1. Illustration of the resultant segments and the differences from the ground truth at Slice 1. (a) The segment in the ROI using traditional FBP and global data—We call it global computed tomography (GCT). (b) The difference between the segment of (a) and the ground truth. (c) The segment in the ROI using scaling function and local data. (d) The difference between the segment of (c) and the ground truth. (e) The segment in the ROI using traditional FBP and local data. (f) The difference between the segment of (e) and the ground truth. The ground truth mentioned here is an image formed based on measured data, illustrated in Fig. 13.7.

from the physical measurements of the target and using the planar projection of the tubular hole structure—ellipses with semimajor axes of 13.2 ± 0.6 mm and semiminor axes of 10 ± 0.8 mm along with the rotation angle of $8 \pm 2^{\circ}$.

Fig. 13.8 illustrated the segmentation processes regarding the planar projection of the tubular hole structure. It should be noted that in the error ratio calculation, what we are concerned with is the difference of the number of pixels between the ground truth and reconstructed segments at each layer. For uniformity, we try to match the reconstructed segments with the ground truth. That is, the position of the centroid of the ground truth is always moved to guarantee an overlapping with the centroid of each reconstructed segment at every segment height. This procedure can be justified via minimized the error between the ground truth and reconstructed segments.

13.4 Reconstruction Results

Four slices are analysed for this target sample: (i) the first slice (Slice 1) is reconstructed with a circular Region of Exposure (ROE) with diameter of 189 pixels and a coincident Region of Interest (ROI) of diameter 110 pixels, both centred in the image; (ii) the second slice (Slice 2) has an ROE diameter of 230 pixels and an ROI diameter of 195 pixels, off-centred from the image centre for ramp filter application onto the image, but an ROI diameter of only 160 pixels, also off-centred from the image centre for wavelet based reconstruction; (iii) the third slice (Slice 3) with an ROE diameter of 209 pixels and an ROI diameter of 150 pixels, off-centred from the image centre; (iv) the fourth slice (Slice 4) with an ROE diameter of 196 pixels and an ROI diameter of 140 pixels, off-centred from the image centre. Each of data set has a pixel interval of 0.5 mm.

13.4.1 Slice One

The reconstructed images at an off-center area with a radius of 95 pixels using the current local reconstruction algorithms are observed. As described in Subsubsection 13.3.2, a 250×250 pixel image of the clown's head target is recovered from scaling coefficients using local data, and the BioSpline2.2 biorthogonal basis is used. This wavelet basis is applied to all the slices for wavelet based reconstructions. A full wavelet reconstruction requires computing the inverse wavelet transform from four images reconstructed using separable scaling and wavelet functions. However, it is found experimentally that

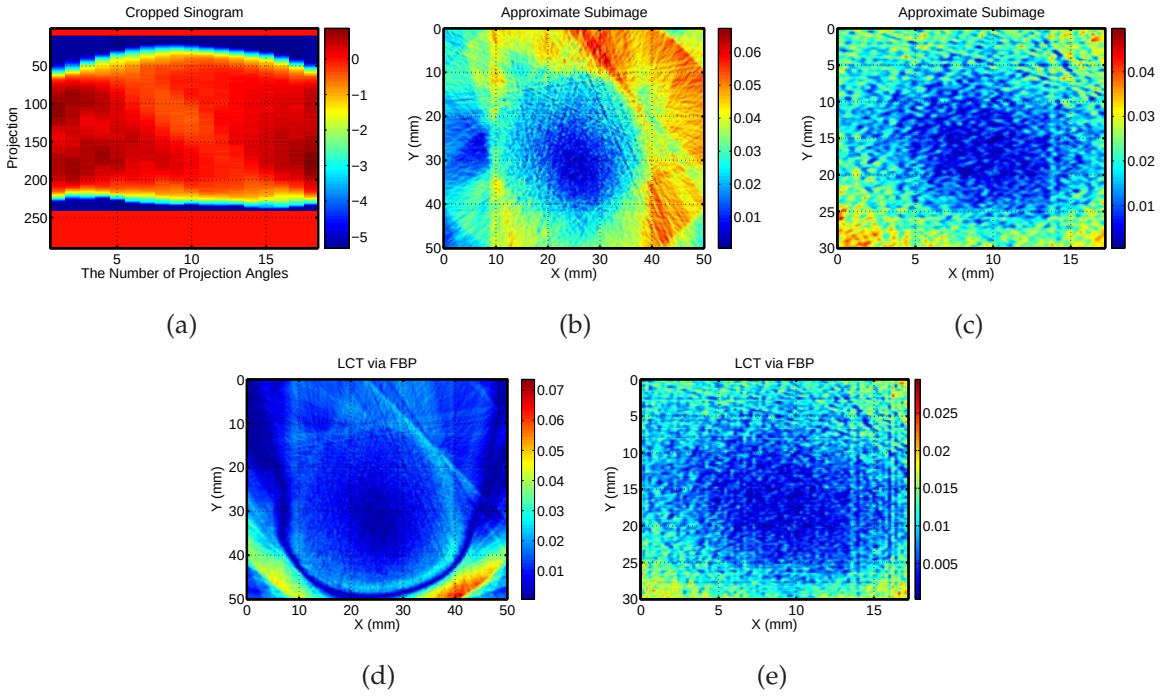


Figure 13.9. Resultant local reconstructions at Slice 2. Illustration of the resultant local reconstructions at Slice 2. (a) Illustration of the truncated projections with 79% of full data at the centered area of interest. (b) Reconstructed image is localised to a region of interest from the scaling function after constant extrapolation. (c) Cropped version of approximate subimage in (b). (d) Reconstructed image is localised to a region of interest via traditional FBP. (e) Cropped version of (d).

reconstructions using only the scaling function provides a good approximation in the region of interest—such results are presented below for illustrative purposes.

The sinogram at Slice 1 with zero padding for nonlocal data is shown in Fig. 13.5. The reconstructed images at an off-center area with a radius of 95 pixels using the current local reconstruction algorithms. A 250×250 pixel image of the clown's head target is recovered from scaling coefficients using local data, shown in Fig. 13.6(a), and Fig. 13.6(b) shows the truncated version of Fig. 13.6(a) showing only the region of interest (ROI). Fig. 13.6(c) shows that the local image with same size is recovered via traditional filtered back projection while Fig. 13.6(d) is the truncated version of Fig. 13.6(c) showing only the ROI. Wavelet and scaling coefficients after back projection are shown in Fig. 13.6(e). Two dimensional inverse wavelet transforms are conducted on the four reconstructed subimages to obtain the image with full package of wavelet based image

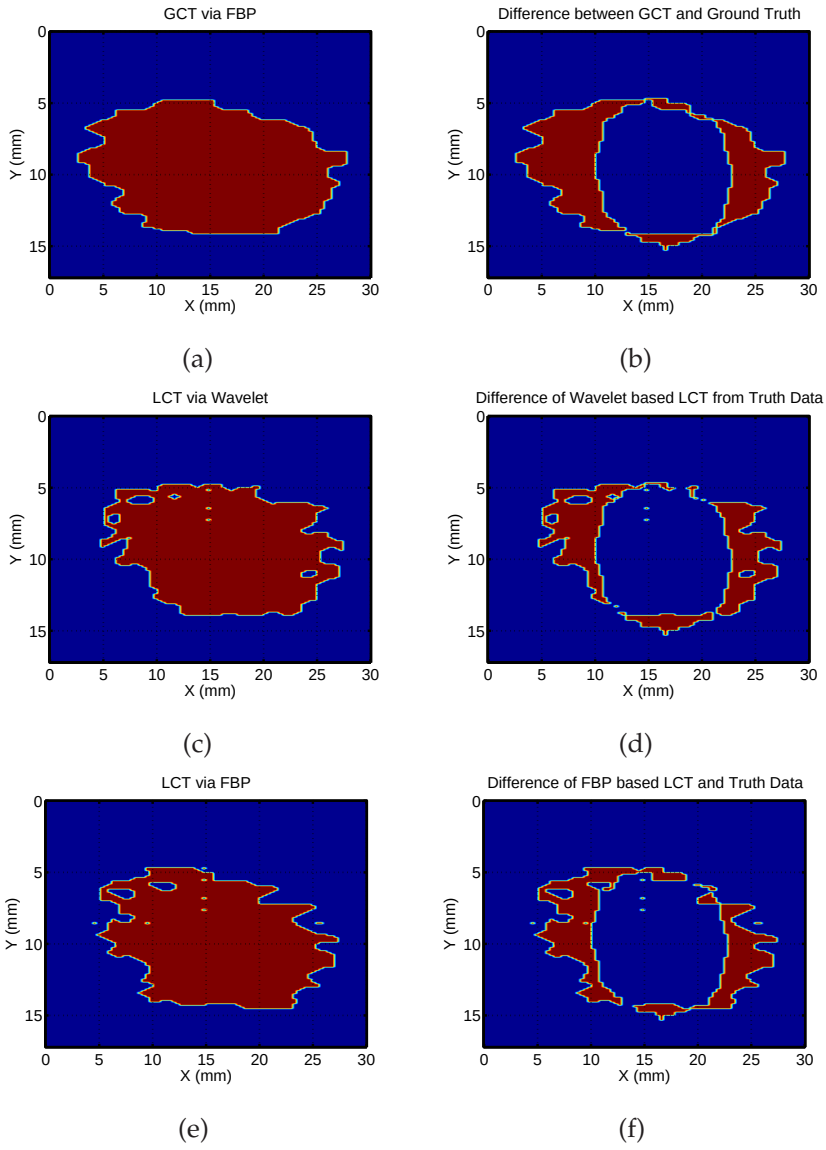


Figure 13.10. Resultant segments and the differences at Slice 2. Illustration of the resultant segments and the difference at Slice 2. (a) The segment in the ROI using traditional FBP and global data (GCT). (b) The difference between the segment of (a) and the ground truth. (c) The segment in the ROI using scaling function and local data. (d) The difference between the segment of (c) and the ground truth. (e) The segment in the ROI using traditional FBP and local data. (f) The difference between the segment of (e) and the ground truth.

reconstruction, shown in Fig. 13.6(f). It is found that in Fig. 13.6, the wavelet approximate subimage of (a) shows higher contrast in the reconstructed intensity than the full inverse wavelet transform reconstruction in (f), which can be illustrated via the associated colour bars. Based on this observation, we only consider the approximate image reconstruction to evaluate the algorithm validation in the current algorithm related to wavelet based reconstruction, because only local information is to be mainly delivered and both the wavelet and scaling ramp filters show essentially same support.

The segment of the reconstructed image via wavelet scaling sequence, Fig. 13.8(c), shows smoother contours with reduced artifacts than the segmented image via the FBP algorithm, Fig. 13.8(e). In addition, the scaling function leads to an LCT with small difference in segment from the ground truth, Fig. 13.8(d), compared to the difference in segment, Fig. 13.8(f), between the traditional local CT and the truth data. But the resultant difference in segmentation from the truth data is a slightly large compared with the difference, Fig. 13.8(b), between traditional reconstructed segment in ROI using global data, Fig. 13.8(a), and the ground truth. The relative error ratio using different algorithms are shown in Table 13.1. Generally, for this slice, the segments show good reconstruction performance whether from the traditional FBP algorithms using global and local data or from the scaling function for LCT, owing to the uniformity in the target sample and the short path length to the hole position.

13.4.2 Slice Two

Fig. 13.9(a)-(e) shows reconstructed images at an off-center area with a radius of 115 pixels using wavelet based local reconstructions and traditional FBP algorithms. It is evident to achieve the same size of the hole feature, the radius of the ROE for local objection reconstruction via FBP is larger than wavelet based reconstruction, both of which result in similar reconstruction error to the corresponding global reconstruction (see Fig. 13.10). Additionally, applying FBP-based LCT to this slice leads to significant ringing problems, but the wavelet ramp filter performs much better in this regard.

Fig. 13.10(a)-(f) shows various reconstruction segments in the ROI and their differences from the ground truth. It is found that localised reconstructions from both the scaling function and the traditional FBP result in reduced reconstruction errors, compared to a global FBP algorithm. However, we have to recognise that the greater scattering from

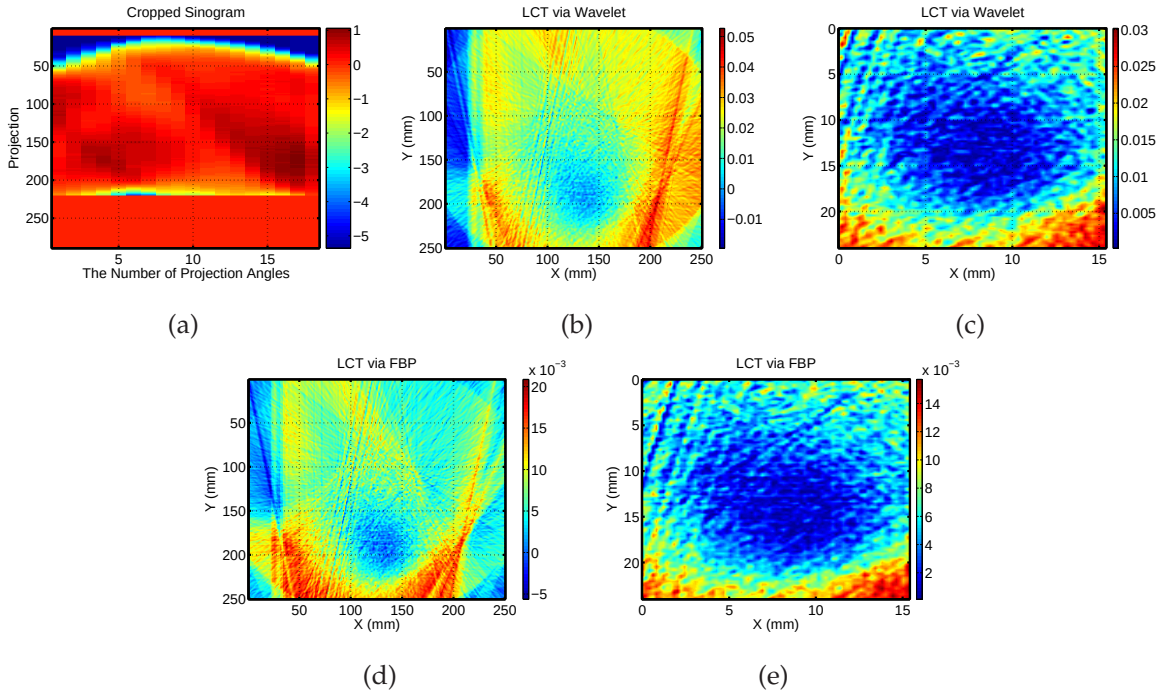


Figure 13.11. Resultant local reconstructions at Slice 3. Illustration of the resultant local reconstructions at Slice 3. (a) Illustration of the truncated projections with 72% of full data at the center area of interest. (b) Reconstructed image is localised to a region of interest from the scaling function after constant extrapolation. (c) Cropped version of approximate subimage of (b). (d) Reconstructed image is localised to a region of interest via traditional FBP. (e) Cropped version of (d).

the irregular shape of the target sample at this height yields more gaps after thresholding; this in turn leads to greater difficulty in the post-processing to bridge the gaps to obtain relatively smooth segments.

13.4.3 Slice Three

Fig. 13.11(a)-(e) shows a series of reconstructed images in an off-centered region of Slice 3 with a radius of 105 pixels and 72% of full projection data. For this slice, the scattering tends to be weakened due to stronger absorption caused by the longer distance between the hole position and the target boundary. As a result, compared to the image reconstruction at Slice 2, Slice 3 shows improved reconstructed accuracy as determined by the difference between the LCT and the ground truth; these are illustrated in Fig. 13.12. In comparison with Slice 2's segmentation performance, the segments

from Slice 3 have smoother boundaries and also reduced misclassification and artifacts due to the larger absorption counteracting scattering effects. Meanwhile, the local reconstruction via scaling function shows improved LCT segments in both smoothness and classification accuracy.

13.4.4 Slice Four

Fig. 13.13(a)-(e) shows, at Slice 4, the reconstructed images at off-center areas and with a radius of 197 pixels and 74% of full projection data.

For this slice, the exterior boundary contour of the target tends to be much smoother than that for Slices 2 and 3. Scaling function reconstruction again shows the better smoothing boundary and reduced misclassification, with stronger intensity contrast in the reconstructed image than the traditional FBP. The local reconstructed segments in Fig. 13.14(a)-(f) show much stronger absorption than scattering since the longer distance from the target boundary area to the hole position.

13.4.5 Segment Evaluation

In order to evaluate the segmentation performance of our experiment, we exploit the fact that the internal structure, the hole, is linear with the same diameter from the top to end. This implies that the segmented hole should have the same dimensions in all slices, with only its centre displaced at various heights. To obtain this measurement, we compute the centroid of the extracted hole for each height, L1-L5, and the resultant x and y locations are recorded in Table 13.2. The slopes computed using centroids from any of two successive slices are recorded in Table 13.3 below, along with the results obtained from our three reconstruction algorithms: traditional global reconstruction, scaling function used for LCT and traditional local reconstruction via FBP. For all the Tables listed below, L1—L4 refer to Slices 1–4, respectively.

Table 13.1 shows the radii of the ROE and ROI at different heights via applying scaling function and FBP algorithms for reconstruction. It is found that, at the height of Slice 1, the radii of the ROE and ROI are the smallest among all the measured heights, owing to the smaller optical distortion, but it also required the greatest value of (ROE-ROI). At the height of Slice 2, the biggest radii of the ROE and ROI are applied to LCT, though,

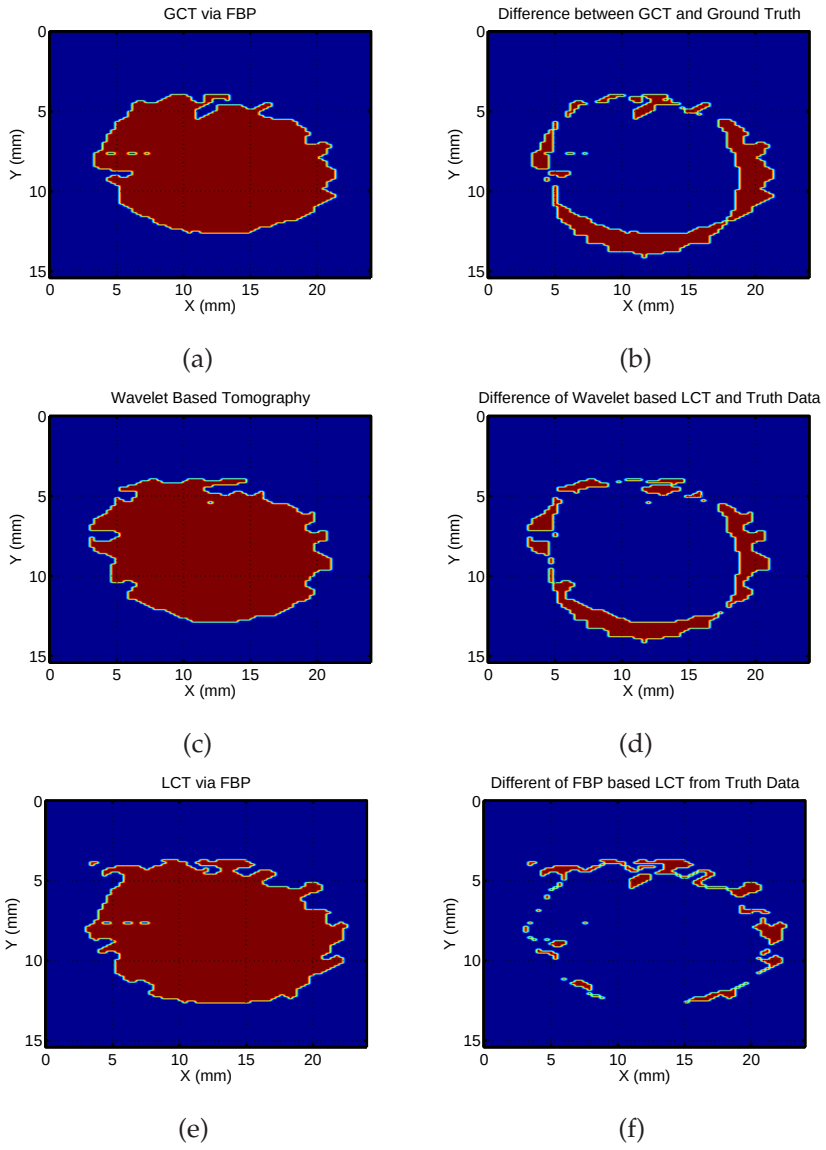


Figure 13.12. Resultant segments and the differences at Slice 3. Illustration of the resultant segments and the differences at Slice 3. (a) The segment in the ROI using traditional FBP and global data. (b) The difference between the segment of (a) and the ground truth. (c) The segment in the ROI using scaling function and local data. (d) The difference between the segment of (c) and the ground truth. (e) The segment in the ROI using traditional FBP and local data. (f) The difference between the segment of (e) and the ground truth.

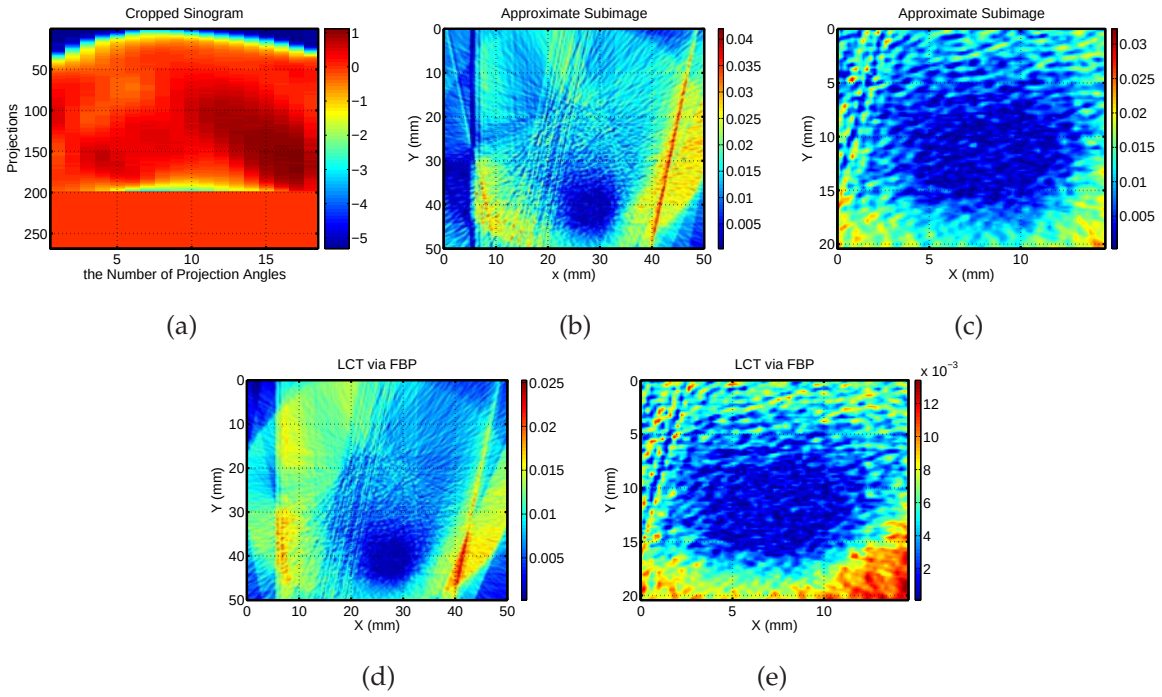


Figure 13.13. Resultant local reconstructions at Slice 4. Illustration of the resultant local reconstruction at Slice 4. (a) Illustration of the truncated projections with 74% of full data at the center area of interest. (b) Reconstructed image is localised to a region of interest from the scaling function after constant extrapolation. (c) Cropped version of approximate subimage of (b). (d) Reconstructed image is localised to a region of interest via traditional FBP. (e) Cropped version of (d).

Table 13.1. Table of the radii of the ROE and ROI. The size of both the ROE and the ROI at four target heights and via the three different reconstructed algorithms, in units of pixel count. The values of ROE and ROI are selected iteratively till the reconstructed hole segment with minimum error in size compared to the 'true' target.

Algorithms	Radius	L1	L2	L3	L4
scaling function LCT	ROE	95	115	105	99
	ROI	55	80	75	70
LCT via FBP	ROE	95	115	105	99
	ROI	55	99	75	70

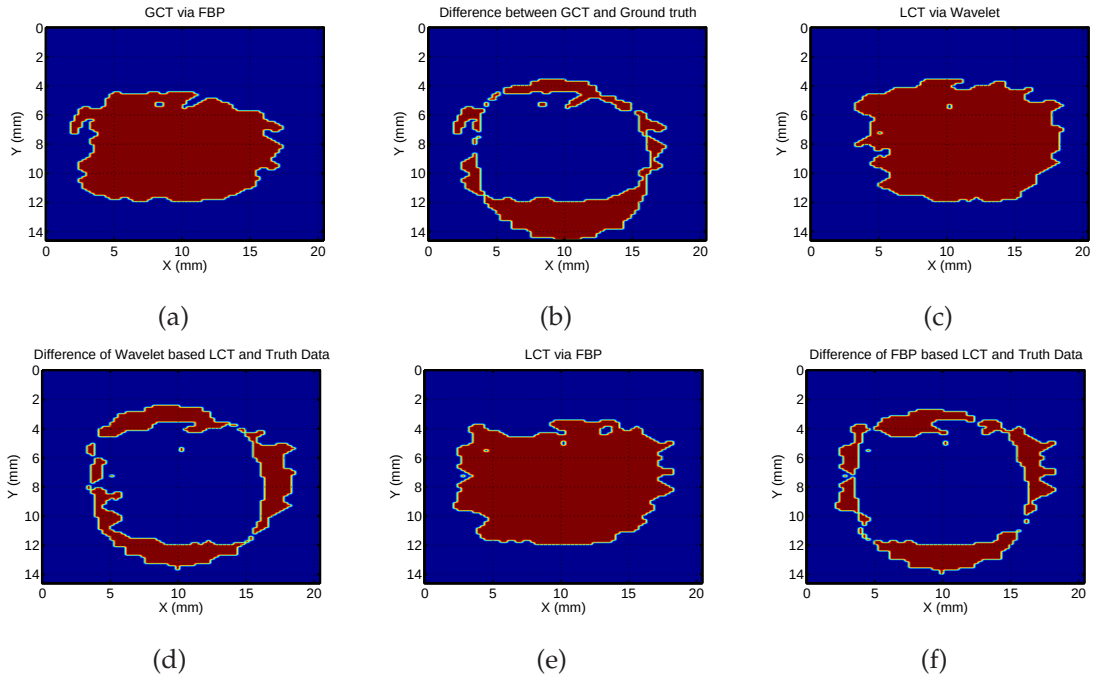


Figure 13.14. Resultant segments and the differences at Slice 4.. Illustration of the resultant segments and the difference at Slice 4. (a) The segment in the ROI using traditional FBP and global data. (b) The difference between the segment in (a) and the truth data. (c) The segment in the ROI using scaling function and local data. (d) The difference between the segment in (c) and the truth data. (e) The segment in the ROI using traditional FBP and local data. (f) The difference between the segment in (e) and the ground truth.

the scaling function for the reconstruction needs a smaller radius of ROI than for FBP based reconstruction. In other words, it is possible for wavelets to acquire the same quality in LCT reconstruction using a smaller exposure area. With the reduced optical distortion and improved reconstruction accuracy at the heights of Slices 3 and 4, there are reduced radii of the ROE and ROI to be employed for the local reconstruction of the similar size of hole cross-section. In summary, less scattering means requiring smaller areas of ROE to recover the same local ROI image, and *vice versa*.

Table 13.2 is the centroid coordinates of the extracted hole cross-section at each height, which are used to calculate slopes at different target heights.

Table 13.2. Illustration of the centroid coordinates. The centroid coordinates of the extracted hole for each height, with units of pixel count.

Algorithms	Coordinates	L1	L2	L3	L4
GCT via FBP	x	134	157	185	207
	y	118	124	130	136
scaling function LCT	x	135	156	184	198
	y	116	126	132	138
LCT via FBP	x	136	154	191	202
	y	121	130	134	141

The slope calculated is a line segment that connects two centroids from any two reconstructed segmented images. For convenience, we call it the reconstructed slope. The error in the slope is calculated via the comparison of the value between the reconstructed slopes and the slope of two 'true' centroids of the hole planar projections at two different measurement heights.

For the local reconstruction, the ROI only contains the hole area. According to the linear structure of tubular hole, the angled projections of the hole at various heights onto x-y coordinate plane has the same x and y coordinates, the result of which is an overlapping with the origin, with slope of 0. It is different from real hole structure with slope of $\tan 43^\circ$, as we only consider the cross-section of the hole with limitation to local hole area, and only pixel measurements is involved. For the reconstructed hole segments at different heights, due to only local data involved for reconstruction, this causes difficulty in accurate recovery of an image. It turns out to be the changed positions of reconstructed centroids at the different hole segments, therefore producing altered reconstructed slopes. Subtracting the true slope of 0 from the reconstructed slope, the error in slope can be viewed as the slope of the line that passes through the centroids of any two reconstructed segments. The average slope is calculated via summing up the slopes for any two different reconstructed segments, which then is divided by the number of collected slopes. The average error in slope across all the reconstructed segments is calculated.

The average slope obtained from scaling function-only local reconstructions is 0.3355, with reduced average error of 0.0112 compared to FBP based local reconstruction. The slope error of the latter is 0.3467. Since the amount of the image data used for the local reconstruction is not complete, reconstruction performance should be inferior to FBP algorithm using global data. The slope calculated using FBP algorithm on global data

Table 13.3. Illustration of error ratio. The error ratio from each of the four heights and each of the three reconstructed methods (It is simply defined as the number of misclassified pixels/total number of pixels).

Applied algorithms	L1	L2	L3	L4
GCT via FBP	0.1380	0.7112	0.3042	-0.3761
approx. LCT	0.1824	0.5306	0.2664	-0.3532
LCT via FBP	0.2136	0.5350	0.3824	-0.3552

yields an average value of 0.2450, which obviously means full reconstruction suffers less distortion than the remaining two local methods.

Compared with the ground truth of a slope of 0, the global reconstruction causes the smallest error in slope among the current reconstruction algorithms. Among the local data algorithms, local reconstruction using scaling function achieves better recovery of the target region than direct FBP on the same local data.

The optical properties are indicated in Table 13.3, via recording the calculated error ratio with respect to the reconstructed segment images on various measurement heights for different reconstruction schemes. Scattering occurs primarily with forward propagation, which involves absorption. Since the only measurable quantity is the intensity of light in the CW THz QCL imaging, and, due to multiple scattering, strong path-length dispersion occurs, which results in a loss of localisation and resolution. The scattering effect affects the image reconstruction, and it turns out that many background pixels are misclassified as hole pixels. The result is that the reconstructed target segment has the size greater than the ground truth. Whereas, the absorption effect results in weakened intensity around the reconstructed boundaries of an object, the result of which is the size or the amount of pixels of a reconstructed object less than the ground truth. The combined performance of absorption and scattering is such that when scattering effect is obvious compared to absorption, the resultant segment tends to have more reconstructed pixels than ground truth; when absorption effect is stronger than scattering, the resultant segment tends to have less reconstructed pixels than the ground truth. Since the current tomographic reconstruction is scattering dependent, generally, the difference between the number of reconstructed pixels and the ground truth is larger than zero, in addition to the layer, where there is obvious absorption effect.

According to Table 13.3, the error ratio is the smallest at Slice 1 and the largest at Slice 2, due to the difference in scattering caused by the variety in shape of external boundary of the target sample at the different heights. The error ratio in Slice 3 is reduced compared to Slice 2 since absorption counteracts the effect of the scattering; the error ratio in Slice 4 is negative because the effect of optical absorption dominates optical scattering.

13.5 Chapter summary

We apply the wavelet based algorithm to reconstruct wavelet and scaling coefficients of a target object image from its Radon transform, obtained experimentally using terahertz signals generated by a QCL. The algorithm is based on the observation that for some wavelet bases, with sufficient zero moments, the scaling and wavelet functions have essentially the same support after ramp filtering. Using experimental data obtained on a 3D target with internal structure, we have shown that wavelet based reconstructions offer robust reconstruction performance in the local reconstructed shape of the target. Segmentation is used in post-processing to make measurements on the target's internal structure, and the results are analysed to determine the effects of the optical distortion produced by varying exterior boundary at the different cross-sections. A reconstruction utilising only the scaling function shows a slight structure distortion in the reconstruction of a 3D structure compared to global CT, due to the use of less data, but superior to traditional FBP algorithm using local data. For the future, a THz spiral CT is an open area for THz tomographic reconstruction. It is promising in achieving computed tomography on a voxel base, as in this case, not only shape information, but structure information about a target object in the local area of interest can be recovered.

Chapter 14

Conclusions and Future Work

THIS Chapter draws together the conclusions from the work described in this Thesis, and proposed directions for further research. The Thesis is composed of two parts. One is for T-ray signal and pattern recognition, and the other is for local terahertz computed tomographic reconstruction. A classification framework was developed for material identification based on THz spectroscopic data. Several different feature extraction algorithms are explored with application on various cases drawn from promising application fields. Combined, these tools pave the way for the development of 3D THz inspection systems with broad applicability. Local terahertz computed tomography has potential in achieving a reduction in the number of THz measurements, in order to improve ability of a THz system to rapidly image. This work solves the problem based on wavelet theory and a novel modification of the inverse Radon transform.

14.1 Introduction

Section 14.2 summarises the research conducted in this Thesis and describes the major conclusions and novel contributions of this work. Section 14.3 then highlights a number of remaining open questions identified in the course of this research and details research areas which form the logical next steps to build upon the contributions of this Thesis.

14.2 Thesis summary

This Thesis is a logical combination of a number of areas, aiming to support two topics: terahertz signal and pattern recognition as well as THz global and local computed tomographic reconstructions. For the THz identification system, the theory scheme mainly concerns a number of key areas of a classification framework, and the relative cases for the validation of different identification strategies via applying THz pulsed imaging systems. For the THz CT reconstruction, we review the methodology of back projection algorithms, and address a segmentation strategy from the perspective of an innovative approach, via a modification of 2D wavelet transforms. THz local coherent tomography is achieved via the modification of a wavelet and scaling ramp filter, in conjunction with the Radon transform. The THz measurements for validating these schemes are obtained for both a THz pulsed imaging system and a CW THz QCL. The Thesis conclusions are presented in the following categories: (i) THz radiation and its sources and detectors, (ii) THz imaging modes, and (iii) THz imaging analysis; methodology regarding THz pattern recognition; case studies for THz identification systems; THz global computed tomography; wavelet based segmentation; pulsed THz local computed tomography; CW THz local computed tomography.

14.2.1 Summary of THz radiation and its sources and detectors

Chapter 1 introduced the T-ray band, spanning frequencies from 0.1 to 10 THz, lying between millimetre waves (mm-waves) and the far-infrared. In the past this has been referred to as the THz Gap due to the lack of sources and detectors with sufficient power and sensitivity. The Chapter addressed the importance of the T-ray band, a theme developed in subsequent Chapters. The gradual development of science and technology in the T-ray region promises to open up new applications, in the same way

as other regions of the electromagnetic spectrum are richly populated with commercial devices and systems.

The T-ray spectrometer in its various forms is described in detail in Chapter 2, including its history, the sources and detectors available, and the tradeoffs in designing and operating these systems. The THz laser sources reviewed in Chapter 2, based on different principles for THz pulsed generation, include Ti:sapphire-based lasers and free-electron lasers. Terahertz quantum cascade lasers are also described as a significant technique providing narrow band sources for THz CW imaging.

14.2.2 Summary of THz imaging modes

THz imaging modes are reviewed in Chapter 3 in three main sections: (i) operation within versus below diffraction-limitations; (ii) pulsed versus continuous imaging techniques; and (iii) time-of-flight imaging and tomographic reconstruction algorithms. The main techniques are transmission and reflection spectroscopy, with both ultra fast laser and THz quantum cascade laser sources.

The main advantages of pulsed T-ray systems over CW T-ray techniques are (i) their high sensitivity, and therefore signal-to-noise ratio, (ii) their broad bandwidth, simultaneously generating all T-ray frequencies from 0.1 to 10 THz depending on the system, and (iii) their ability of THz radiation to resolve the materials absorption from scattering.

The review of the current state-of-the-art in THz imaging surveys the current opportunities and limitations in the THz imaging field. Each imaging mode is a platform on which to achieve further signal and pattern recognition using THz radiation. It then motivates the development of a local tomographic imaging system, which is then described in detail in this Thesis and validated by pulsed THz measurements and THz QCL tomographic imaging.

14.2.3 Summary of THz imaging analysis

Terahertz image analysis is reviewed in Chapter 4. The processing techniques are used to analyse the THz interaction with a target specimen, which enhance an image. The techniques reviewed in this Chapter involve (i) *k*-means classification; (ii) the separation of the component spatial patterns; (iii) the Euclidean distance classifier; (iv) a

support vector machine for classification of the terahertz relevant frequencies. A number of target samples are considered, including DNA, RNA, and retinal isomers, to biological tissues, cells, such as, tooth, brain tissues, skin and breast cancel cells; from tablet coating to water, liquid, etc. Terahertz spectroscopy for biomedical signals is analysed in the time domain, the frequency domain, the time-frequency domain, and the wavelet domain. The review of the state-of-the-art in THz imaging analysis shows the potential in the development of THz identification systems, and motivates the THz pattern recognition research, which forms one part of this Thesis.

14.2.4 THz pattern recognition

We adopt pattern recognition algorithms in order to address the goal of differentiating heterogenous layers within a target, via interaction between terahertz radiation and the target medium. There are a series of potential applications for THz pattern recognition, such as security screening, quality control, biomedical diagnosis, etc.

Chapter 5 develops a pattern recognition framework for the identification of materials vial pulsed THz images. Chapter 5 is composed of three parts: data acquisition through THz imaging, data preprocessing, and decision classification. The preprocessing procedure is realised through two important steps including data preprocessing and feature extraction. The exploration of feature extraction methods is a focus of this Thesis. The extracted features then are input to numerical different classifiers.

14.2.5 Wavelet and preprocessing

Wavelet transforms are popular techniques suited to the analysis of signals that are localised in time. They well suit to signals with sudden and unpredictable changes that often carry the most interesting information. In this Thesis, wavelet transforms provide important analytical techniques for most of the work carried out in this Thesis.

Chapter 6 presents a review of the theory and practical techniques of applying wavelets to signal and image analysis. The focus of our review on wavelets is on their multiresolution analysis and computational properties, with a link to Mallat's pyramid algorithms. These are the primary reasons for choosing wavelets as the analytical tool. Discrete wavelet transforms and wavelet packet transforms are mainly represented by this Chapter.

Data preprocessing attempts to isolate material information present in the THz waveforms from systematic and random noise sources present in the data, which takes an important role in THz pattern recognition. Wavelet soft threshold denoising is also demonstrated in this Chapter for this purpose.

14.2.6 Feature extraction and selection

Feature extraction aims to extract key features from a pattern. Mathematically speaking, this leads to a dimensionality reduction for the THz data set. Chapter 7 describes five various feature extraction algorithms for the resolution of different THz pattern recognition problems.

A novel technique involving the use of Auto Regressive (AR) and Auto Regressive Moving Average (ARMA) models on the wavelet transforms of T-ray pulse responses is applied to feature extraction methods for classification purposes. The correlation and Durbin's methods are viewed as optimal in the calculation of the AR and ARMA model parameters. A prediction error minimization criteria is applied to achieve accurate signal evaluation based on local optimization algorithms.

The subspace identification algorithms applied in this Thesis accord with the work carried out by Galvão *et al.* (2005). The identification system determines state sequences by projecting the input and output data sets. Since no explicit optimization criteria is involved, subspace methods are found to be fast and numerically stable, but less accurate.

A subband system identification framework is applied to explore the spectral partitioning flexibility of wavelet packets with use of THz radiation. It is an extension of the work performed by Paiva and Galvão (2006). The best wavelet packet tree is selected by calculating the cost of each leaf node in a wavelet packet tree. The number of leaf nodes in the best tree is adjusted by a penalty factor, calculated via minimising generalized cross-validation (GCV) function.

System identification via optimized Mertz apodization functions is also discussed. The design of Mertz asymmetric triangular window is a trade-off of minimizing sidelobes of the calculated spectrum. It turns out to be an increasing main lobe width, which leads to a reduction in frequency resolution. Tailored asymmetric apodization functions are also necessary for the far more asymmetric THz transient waveforms.

14.2.7 Pattern classification

The final step for pattern recognition is pattern classification. In this stage, the extracted and conditioned features are grouped together into different classes to produce the final identified pattern. There are many pattern classification techniques in the pattern recognition literature. Chapter 8 describes a few common ones. The Mahalanobis distance classifier is a quadratic classifier. It is a very useful way of determining similarity of an unknown sample set to a known one. Other examples of classifiers or classification scheme described in this Chapter are: the Euclidean discrimination matrix and Support Vector Machines (SVM).

A supervised algorithm is adopted to provide a class label for each pattern. Cross-validation is a statistical technique widely used for estimating generalization error, which consists of the holdout method, *K*-fold cross-validation, and leave-one-out cross-validation. In addition, measure functions (or matrices) are as an alternative way of selecting and evaluating learned classifiers.

14.2.8 THz pattern recognition experiments

With each element in our THz pattern recognition system having been discussed, the system is tested through extensive experiments. Chapter 9 contains the detailed results from the experiments performed.

The powder spectroscopy study demonstrated that Rayleigh scattering is a critical concern in THz propagation through powdered substances. The aims of the current research is to attain better understanding of THz scattering effects (Pearce and Mittleman 2001). The classification of materials independent of density, particle size and thickness is a difficult problem and would benefit from advanced models of THz propagation in random media. We design a number of models to explore the powder classification performance with various thickness, density and particle size. (i) Frequency feature selection via a simple iterative algorithm and the Mahalanobis classifier allows good classification performance with identification accuracy reaching 98.6%. (ii) THz feature extraction via ARMA model of wavelet subband coefficients with Mahalanobis distance as a classifier, permits a classification accuracy reaching 98.0%. This model avoids to select features from over 100 sampled frequencies. From a computational complexity point of view, the algorithm exploiting the ARMA model of wavelet subband coefficients is advantageous. (iii) A multiclass classification is investigated via

support vector machines. THz frequency features are selected via frequency orientation component, which are then randomly selected as training sets and testing sets are input to the support vector machines. Under this pattern identification system, we consider several powder samples at a thickness of 2 mm. The initiate exploration of the application associated with support vector machines may prove beneficial to further development of THz identification systems.

Classification of lactose and mandelic acid THz spectra is also investigated. The goal of this work is to demonstrate efficient and robust classification algorithms that could be adopted by the biomedical and pharmaceutical communities. Three different system identification algorithms are applied to achieve THz signal evaluation with gradual degradation of responses by increasing the noise level. A Mertz apodization window for ratioing the sample performs good in smoothness and in the identification with respect to the absorption bands. Subspace identification system and wavelet packet techniques, however, show superior performances of different sample discrimination with respect to different noise levels. A wavelet packet technique is shown to be particularly suited to the identification of systems with several spectral resonance features.

The cancerous cell case study aims to complement recent work on the detection of basal cell carcinoma tissue using THz spectroscopy. To isolate the cellular response of cancerous cells from the multitude of complicating factors encountered in *in vivo* studies, an *in vitro* approach is adopted. Normal human bone cells and osteosarcoma cells are cultured in polyethylene flasks. Once the cultures are imaged using a THz imaging system and the spectra analysed under the developed classification framework—wavelet packet SURE soft thresholding for preprocessing; AR modeling over wavelet packet subband; Mahalanobis classification scheme. Once again, the results are promising and form a foundation for future in-depth studies.

Application of T-ray techniques to identify the binding state of DNA is an especially interesting topic. It is expected that the THz radiation is capable of differentiating the different types of DNA, though there is lack of characteristic absorption features in the T-ray region. The recognition of the frequency domain data from two different RNA polymer strands, polyadenylic acid (poly-A) and polycytidylic acid (poly-C) is achieved with a support vector machine. The classification performance is comparative with unsupervised learning algorithms, which provides the technology required for the further proliferation of THz-transient spectrometers.

In all these case studies, high classification accuracies are demonstrated. These studies highlight the performance of the pattern recognition strategies developed in Chapter 5 to Chapter 8, but they are, by nature, preliminary. Rather than attempting to perform comprehensive studies in these application areas, these studies sought to highlight the potential of THz inspection systems and to provide a basis for future work.

14.2.9 CT and THz CT

Computed tomography and terahertz computed tomography are overviewed in Chapter 10. T-ray computed tomography provides spectroscopic 3D images with high fidelity. The Rytov approximation to the Helmholtz equation is adopted to derive a series of experimental conditions under which the filtered back projection algorithm (similar to X-ray CT applications) are utilised to reconstruct targets based on the measured THz field.

THz CT has slow acquisition time, but very high SNR and image quality. It is a novel extension of terahertz time-domain spectroscopy with numerous potential applications. We extract the frequency dependent refractive index of a 3D target thereby providing spectroscopic images of weakly scattering objects. T-ray CT provides the refractive index of the sample without requiring *a priori* knowledge of the sample thickness and allows the internal structure of objects to be revealed.

An interpolated cross-correlation algorithm is applied to estimate the phase shift of the THz pulse after transmission through low-dispersion targets. The time domain algorithm as an alternative approach to achieve Radon transform images can be utilised as part of the reconstruction algorithm.

14.2.10 2D wavelet-based segmentation by fusion in 3D THz CT

Chapter 11 investigates a novel segmentation technique for terahertz (T-ray) computed tomographic (CT) imaging. A set of linear image fusion and novel wavelet scale correlation segmentation techniques is adopted to achieve material discrimination within a three dimensional (3D) object. The methods are applied to a T-ray CT image data set taken from a plastic vial containing a plastic tube. This setup images a simple nested structure, with potential to achieve T-ray pulsed signal classification of heterogeneous layers. The wavelet based fusion scheme does not require the calculation of a single threshold and there is only a segmentation parameter to be adjusted.

14.2.11 Wavelet-based terahertz coherent local tomography

Chapter 12 develops a local computed tomography algorithm according to global THz computed tomography systems, which represents the culmination of this research on THz imaging systems. A wavelet-based algorithm is applied to reconstruct THz-CT images with a significant reduction in the required measurements when the region of interest is small. Instead of inverting the interior Radon transform, the approach uses modified wavelet and scaling ramp filters along with the traditional back projection algorithm to obtain the resultant reconstruction.

The algorithm recovers an approximation of the region of interest from terahertz measurements within its vicinity, and thus improves the feasibility of using terahertz imaging to detect defects in solid materials and diagnose disease states for clinical practice. Several test targets are fashioned from polystyrene and used to characterise the system.

14.2.12 Local computed tomography using a THz QCL

Quantum cascade lasers (QCLs) are semiconductor injection lasers, with the emission wavelength controlled by the thickness of the MQW. Intrinsic high-power capabilities of the lasers make possible the cascading process, with several tens of photons generated in superlattice structures.

Chapter 13 investigates local reconstruction of the region-of-interest (ROI) from a 3D terahertz imaging obtained via a quantum cascade laser (QCL). The optical properties of THz QCLs in ROI are also studied when interacting with materials having complex contours (a clown's head with a hole inside). This Segmentation algorithms are applied on the local reconstructed images with low contrast to observe the QCL imaging ability.

It is found that 3D local reconstructions of the target (hole) using a QCL take on a number of different shapes since the various contours of the target physically distort the measured optical parameters of the object. Local computed tomography via wavelet is suitable for the image reconstruction in terahertz frequency range with lower image quality, which results in lower misclassification after segmentation, than traditional FBP algorithms. A wavelet scaling function allows to deliver local information for reconstruction, which shows a slight structure distortion in the reconstruction of a 3D structure compared to GCT, due to the use of local data, but better than the traditional FBP algorithm using local data.

14.3 Future Directions

With any rapidly developing technology there are a vast number of open questions and promising future research problems. THz inspection systems are no different. This Section surveys the scope of the future work in this area and particularly highlights promising extensions of the work presented in this Thesis.

14.3.1 Curvelet transforms

One focus of this Thesis is on the wavelet transform aspects of THz responses. However, more advanced mathematical models, such as the ridgelet and curvelet transforms, have been proposed as alternatives to wavelet representation of image data. Conceptually, the curvelet transform is a multiscale pyramid with many directions and positions at each length scale, and needle-shaped elements at fine scales. It should be very valuable to an image system. The main benefits reported by Candès *et al.* (2006) are (i) enough frequency localization and spatial localization; (ii) the statistical optimality of the curvelet shrinkage tends to well recover edges on ill-posed inverse problems; (iii) curvelets also have special microlocal features which make them especially adapted to address certain reconstruction problems with missing data; (iv) curvelets especially perform well in noise threshold, which was detailed represented by Starck *et al.* (2002).

It is interesting to examine the effects of implementing curvelets into the THz imaging system. In particular, THz imaging with different noise level, such as cancerous cell images. Meanwhile, THz computed tomographic reconstructions via curvelets are also attractive, since the improved time-frequency localisation feature.

It will be a challenge to identify a suitable measure or objective from the THz parameters, since sharp spectral features are not observed at room temperature using THz radiation. Effective noise reduction from curvelets will lead to good identification performance between different types of biomedical materials using T-rays. In addition, this Thesis touches upon classification experiments with frequency domain transforms and some types of statistics that can be extracted from wavelet transform coefficients as the primary features for THz radiation separation. How about curvelets? It remains to be investigated, which may lead to improvements in classification performance of THz responses.

14.3.2 Tomographic reconstruction

Tomographic reconstruction is another area of focus of this Thesis. There are many techniques regarding computed tomographic processing.

Spiral computed tomography

Terahertz spiral computed tomographic (CT) scanning: is an improved version of terahertz CT imaging technology. Spiral CT (Brink *et al.* 1995, Kalender *et al.* 1990) is an especially advanced technology in the reconstruction of an object in a volume space, which allows continuous increments in measurement heights, with reduced system movement, and centre point alignment of each image. It may be designed and applied for 3D object reconstruction. Local tomography based on THz spiral CT scanning may allow rapid measurements of THz volume data.

The development of spiral computed tomography (SCT), which uses a slip ring CT scanning stage to obtain object rotations coupled with continuous increment of movement, can permit volumetric acquisitions (Kalender *et al.* 1990). This may be a novel and promising experimental contribution to terahertz imaging. During the course of this study, several technical advancements will occur for performing spiral CT scanning, including the exploitation of interpolative and extrapolative techniques.

Statistical reconstruction

THz local CT will be an enabling technology for terahertz functional imaging that allows measurement of the target THz spectrum in the area of interest and in a limited measurement space. Statistical reconstruction allows a stable solution of limited-angle local tomography; a severely ill-posed inverse problem (Rantala *et al.* 2006). It is suitable for an improved inversion since *a priori* information about the unseen target can be efficiently combined with limited image data (Hanson and Wecksung 1983). Bayesian models (Kolehmainen *et al.* 2003, Herman 1980) are applied to achieve statistical reconstructions to reduce the number of projection angles. Wavelet expansion is suggested to reduce the dimension of the computational problem before reconstruction.

Currently, most reconstruction algorithms are based on an ideal deterministic relationship between measurements and the reconstructed object. For local tomography, the limited available data in number and range of viewing projection angles may not be sufficient to specify any reconstruction uniquely. The special aim of Bayesian models,

here, is to achieve statistical reconstruction to reduce the number of projection angles. Feasibility of the statistical methods has been demonstrated by numerical examples using *in vitro* data via X-rays.

Molecular probes for *in vivo* THz fluorescence imaging

Fluorescence-based observation and imaging of living systems is one of the most rapidly evolving fields in clinical and experimental biology today (Graves *et al.* 2004). It is promising for sensitive and accurate THz detection of early stage cancer. Current methods for cancer detection involve the differentiation of the diseased tissue from normal cells via the contrast of different water content. However, this method does not greatly differentiate between cancerous and normal cells, and is a late diagnosis with cancer cells invading other parts of the body. Recent research in molecular imaging has developed functional nanoparticles that are covalently linked to biological molecules such as peptides, proteins, or small-molecule ligands (Alivisatos 2004, Alivisatos 1996). It is helpful to understand THz inspection of cancer via imaging contrast medium.

Fluorescence tomography using T-rays: This work followed original observations that THz has the depth resolution of 5 to 6 mm in moist tissue, and a maximum of 11 mm in dry tissue (Mickan *et al.* 2000). Penetration depth of THz is comparable to OCT (Woodward *et al.* 2003, Fitzgerald *et al.* 2005), but with reduced Rayleigh scattering (Yin *et al.* 2007a). For imaging applications, a single point on the tissue boundary is illuminated and the tissue can be viewed at multi angles via a rotation stage connected with a linear stage on which the sample is mounted. Patterns that T-rays interact with tissue are collected at the different position and different rotation angles around the boundary of the moving sample via a photodetector or a CCD camera. These THz measurements of appropriate fluorescent molecules are then combined in a tomographic scheme, for reconstruction of intrinsic tissue contrast, that is, absorption or refraction. This THz imaging mode with specificity to cellular and sub-cellular processes, will lead to the development of fluorescence molecular THz tomography, a technology directed towards non-invasive quantitative molecular imaging of live tissue.

14.4 Summary of original contributions

The original contributions represented by this work are discussed in Section 1.4. In summary, they include:

- This Thesis explores Multi Resolution Analysis (MRA) techniques, and illustrates that MRA can be applied to (i) denoise terahertz signals by modifying wavelet transform coefficients (discrete wavelet transforms and discrete wavelet packet transforms) for accurate classification (Yin *et al.* 2007b); (ii) decompose terahertz images to extract target segments for image enhancement (Yin *et al.* 2007e); (iii) reconstruct local areas of an image due to the superior time-frequency localisation characteristics of wavelets (Yin *et al.* 2007c, Yin *et al.* 2008a).

- In the terahertz sensing area, the Thesis shows visualised classification of two-class and multiclass terahertz responses via a series of different pattern recognition strategies, highlighting the significance of terahertz spectral information (See Section 9.5) (Yin *et al.* 2007g, Yin *et al.* 2007h).

- The Thesis designs a novel multiresolution based segmentation (Yin *et al.* 2007e) algorithm. Chapter 11 performs initial investigations into the exploitation of (i) the wavelet scale correlation approach to extracting important segments from terahertz CT images; and (ii) the evaluation of terahertz CT segmentation technology from the geometric point of view.

- Statistical models presented in Section 9.4 for cancer cell identification are one of the important contributions of the Thesis on THz pattern recognition. It is discovered that statistical modelling (AR/ARMA) using wavelet coefficients can be applied to extract highly descriptive features, and produce improved classification performance of cancerous cells from normal tissue (Yin *et al.* 2007g). RNA signal identification is investigated using support vector machines. Section 9.5 performs the very first study on support vector machines (SVMs) for classification of 2-class RNA patterns & multiclass terahertz pulses (Yin *et al.* 2007h).

- Terahertz local computed tomography (LCT) is investigated in Chapter 12. The analysis methods for the local reconstruction of terahertz CT data are realised by applying separable wavelet ramp filters for mapping THz measurements in the area of interest in an image. Realisation of terahertz CT by applying wavelet theory is a novel approach to THz imaging (Yin *et al.* 2008a).

- Optical analysis of QCLs is an extension of the current terahertz LCT reconstruction. It is the result of collaboration with researchers at The University of Cambridge. Chapter 13 performs initial investigations into the exploitation of the

various contours of a dielectric target that physically change the optical properties of the object, when terahertz QCLs are used for CW imaging. It also exhibits multiresolution based local CT has improved reconstruction properties over traditional CT reconstruction techniques (Yin *et al.* 2008b).

- In collaboration with researchers at The University of Reading, we realise the identification of lactose and mandelic acid THz spectra using subspace and wavelet packet algorithms, which are overviewed in Section 9.3. It mainly involves the complex insertion loss function of THz transients of lactose, mandelic acid, and dl-mandelic acid via three signal processing algorithms: (a) standard evaluation by ratioing the sample with the background spectra, (b) a subspace identification algorithm, and (c) a novel wavelet packet identification procedure (Yin *et al.* 2007d).

14.5 In closing

Pattern identification and image recognition have been important research topics in computer science. This Thesis realises THz recognition and tomographic reconstruction with novel techniques. This Chapter gives an overview of the major conclusions of this Thesis, and presents a number of recommendations for future work in the field. With continued progress in the field, practical THz computed tomographic image and pattern recognition systems are foreseeable in the near future.