

WAVELET BASED SEGMENT DETECTION AND FEATURE EXTRACTION FOR 3D T-RAY CT PATTERN CLASSIFICATION

X.X. Yin, B.W.-H. Ng, B. Ferguson, S.P. Mickan, D. Abbott

School of Electrical & Electronic Engineering, The University of Adelaide, SA 5005, Australia

ABSTRACT

This paper explores three dimensional (3D) Terahertz (T-rays) computed tomographic (CT) classification based on T-ray functional imaging techniques [1]. The target objects are separated by their refractive indices, which are indicated by the intensity in the images. Segmentation techniques are employed to identify the position of each pixel belonging to the different classes. Wavelet methods are applied to the detected T-ray pulsed responses for feature extraction. A Mahalanobis distance classifier is selected for the final classification task.

This paper presents T-ray CT Classification techniques that allow analysis of measured T-ray transmission image statistics and that automatically identify materials within a heterogeneous structure.

Index Terms— Terahertz computed tomography, Segmentation, feature extraction, Mahalanobis distance classifier

1. INTRODUCTION

'T-rays' is a collective term to describe the part of the electromagnetic spectrum from 0.1 THz to 10 THz. The application of T-rays, especially in the biomedical and security fields, is attractive owing to two intrinsic properties: a non-ionising nature and strong penetration through dry, non-polar and non-metallic materials. Rapid improvements in T-ray detectors and sources make it possible to image objects through optically opaque layers. At present, three dimensional (3D) functional T-ray CT imaging, described in [1], has been developed based on T-ray frequency responses. T-ray CT imaging promises to make vital contribution to a large number of clinical applications, particularly as a means of assisting clinical diagnosis. In this paper, the 3D classification for T-ray CT functional imaging is investigated. A set of linear image fusion and segmentation techniques, including a novel wavelet scale correlation method, are adopted to achieve the discrimination of materials within a 3D object. The methods are applied to four dimensional (4D) T-ray CT imaging datasets of a glass vial containing a plastic tube. This experiment, as realised in Fig 1, 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.

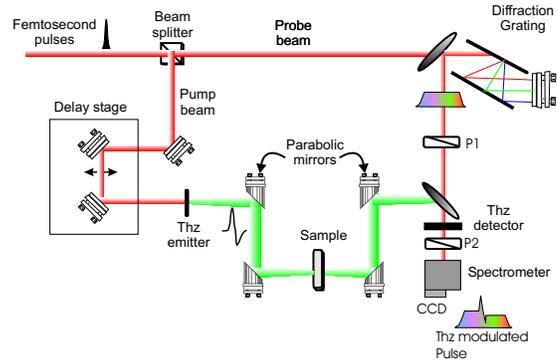


Fig. 1. A chirped probe pulse THz imaging system is demonstrated to realise a 2D T-ray CT imaging. After [1].

In this paper, a wavelet-based classification scheme for a three dimensional object is presented. The scheme's classification accuracy is estimated using a Mahalanobis distance classifier. The performance of the scheme is compared for three different types of wavelets: orthogonal, biorthogonal and non-orthogonal.

This paper contributes to the T-ray and image processing research field by applying image fusion and segmentation techniques to 3D T-ray CT imaging with an aim to achieve the eventual classification of homogeneous materials embedded within 3D objects.

2. THE MOTIVATION

T-ray time domain techniques allow the measurement of T-ray pulsed response in both amplitude and phase (time delay). On the one hand, the property of the material under investigation with heterogeneous layers can be analysed via the complex dielectric function—both amplitude and phase (time delay)—when comparing input and output waveforms. Here, we only demonstrate the output waveforms, neglecting the inputs in order to highlight the differences between two variable output waveforms. As illustrated in Fig. 2, two different pixels are selected at a height equal to 7 mm and at two different ro-

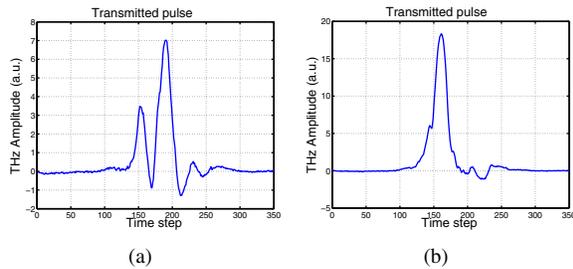


Fig. 2. Two different pixels are selected from the T-ray image data at a height equal to 7 mm and at two different rotation angles. (a) Transmitted T-ray pulse with two obvious peak value modified by passage through target substance—the vial and the tube. (b) Transmitted T-ray pulse modified by passage through target substance. One of the two peaks has been weakened greatly, while the other turns larger.

tation angles. One of transmitted pulses is shown in Fig. 2(a), with two various peak values as a function of time, which affords one dimensional information for depth scan of the target sample structure—the vial and the tube. It exhibits greater contrast between the two target classes and the background. It is convenient to separate target samples corresponding to the different classes via the time slice. However, as shown in Fig. 2(b), it is difficult to divide the transmitted pulse via time slice for three classes: the vial, the tube and the background, due to the lack of obvious variety of peak and valley after the penetration of T-rays through three target objects. This is because, at different rotation angles and for different tested objects, there are different refractive, scattering and absorption coefficients, which result in a great decrease in one of the peak values of the response and a relatively large increase in the other.

A simple consideration for the slice of different target classes is the application of segmentation techniques to T-ray CT functional imaging, which affords an effective approach to define the regions for different classes, i.e. classification of T-ray CT images for functional imaging.

3. THE METHODOLOGY

The current T-ray CT classification scheme consists of several stages: T-ray functional tomographic imaging for segmentation, feature extraction achieved by discrete wavelet transform and the final classification via Mahalanobis distance classifier.

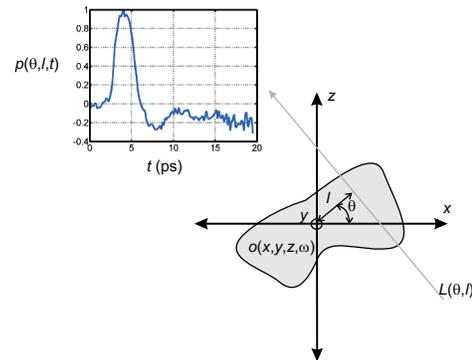


Fig. 3. The coordinate system for T-ray CT. The x , y and z axes come from the standard Cartesian coordinates (y axis is perpendicular to the page). The object is rotated around y axis. The angle θ is from the x axis to the THz beam which propagates through a target. L offsets from the rotation axis — y axis. A typical THz pulse spectrum is shown as a function of time t in the inset. After [1].

3.1. A Brief Introduction to T-ray Functional imaging

Terahertz pulsed imaging (TPI) is achieved by repeating pulsed Terahertz measurements in a 2D raster scan. An ultrafast pulsed laser beam is split into separate probe and pump beams. The path length of the pump beam is modulated by a delay stage, then transmitted through a chopper and enters one of the optical rectification crystals, which acts as a T-ray emitter. The T-rays produced are recollimated and focused onto a sample by a pair of optical lenses. The T-rays emerging from the sample are adjusted again by another pair of optical lenses before being combined with the probe beam. As a result, the T-ray response and the probe beams propagate through the THz detector crystal co-linearly. The detector crystal produces an optical output which is proportional to the T-ray response, and this signal is measured with the use of a photodetector.

Depending on the different peak time of the amplitude varieties, a set of 2-D surface data of an object can be obtained according to the different object angles of the T-ray and the change of the position along the X-axis. To form the 3-D image data, another Z-axis should be built so that a spatial target image of the different cross-sections can be reconstructed.

Fig. 1 shows a 2D chirped pulsed imaging system. An optical probe beam is chirped by a grating pair and allows an electro-optic crystal to sample the THz temporal profile simultaneously.

The coordinates before and after the reconstructed CT image are illustrated in Fig. 3. In the current experiment, a four-dimensional (4D) dataset is acquired with coordinates (θ, l, z, t) , where θ is the projection angle, l is the projection lines' distance from origin, z is the vertical axis and t is the

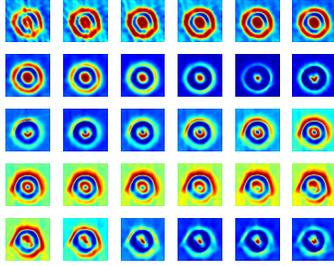


Fig. 4. Illustration of the T-ray CT images at thirty different frequencies with a height equaling to 7 mm and a time step equaling to 0.15 ps.

time. The corresponding set of coordinate axes labelled by (x, y, z, ω) is employed for the reconstruction of the object's optical properties, where ω indicates the frequency, and the x, y and z axes come from the standard Cartesian coordinates (the y axis is perpendicular to the paper). The object is rotated around z axis.

3.2. Functional imaging for segment detection

In principle, T-ray CT extracts the frequency dependent 3D characteristics of a target sample [1]. This requires the measurement of four dimensional datasets to describe the sample. T-ray CT images can be reduced to 2D projections from a 3D volume space, which facilitates the identification of target materials. Briefly, the T-ray Computed Tomography (CT) imaging procedure consists of computing an inverse Radon Transform on the Fourier transform coefficients of the measured signals. The Wiener preprocessing of T-ray pulsed signals plays an important role in reducing noise and achieving better performance of T-ray CT imaging with improved smoothing and resolution. The frequency spectrum of T-ray signals depends on the full THz time domain response, which is measured via a chirped probe pulse THz imaging system [1], shown in Fig. 1 (b). The resolution of the reconstructed image improves with increased frequency, but this is achieved at a cost of decreased signal to noise ratio (SNR).

In this experiment, the imaged object has an internal structure of a plastic tube inserted inside a glass vial. Since the glass is relatively transparent to T-rays, the internal plastic tube is visible. This object is imaged at heights from 5 mm to 9 mm (from bottom) in 1 mm increments. Fig. 3 shows the reconstructed T-ray CT image via Inverse Radon Transform (IRT) at thirty different frequencies at a height of 7 mm. In the experiments, the sample was imaged from bottom to top at five different heights and the rotation angle internal was 10° . Each T-ray CT image consisted of 89×89 pixels, while the number of the samples in each time domain signal was 350 with a time step of 0.15 ps. In order to achieve the T-ray CT

signal classification, the pixels at relevant heights were segmented for each class. A 2D wavelet scale correlation method consisting of two-scale 2D wavelet transforms and convolution of the approximate subimages were employed for the segment extraction.

This procedure of 2D wavelet scale correlation method contains two stages. Firstly, the target cross-sections are assumed to be corrupted by additive white Gaussian noise, which is distributed randomly. The target objects are separated by their absorption coefficients, which are indicated by the grey scale intensity of the images. With the incremental wavelet scale, the local wavelet modular maxima of noise is reduced and the target intensity (energy)—local wavelet modular maxima of the signal is increased in an image. This is determined via the different transmit property of wavelet transform local modular between the signal and noise [2]. After computing 2D DWTs [3] of fused T-ray CT images, the portions for the increased energy can be viewed as the target segments. Therefore, an increased energy with an increase in wavelet scale is used as a cue to extract the target regions.

The Canny edge detector and Otsu's threshold method [4] were used to separate the different target classes. Figures 5(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. Therefore, each pixel is classified as either glass vial, plastic tube or the background. The labelled one dimensional signals in the time domain corresponding to one of the segmented pixels from three class samples are shown in Fig. 6. In this figure, in order to clearly show the differences related to the reconstructed signals of the three classes, the first 60 time steps are applied. It is obvious that the signals from three target samples are well differentiated. Furthermore, the one dimensional T-ray responses represent important frequency ranges for signal interpretation than the two dimensional image statistical data which is reconstructed at one frequency.

3.3. Feature Extraction

Feature extraction is an important step in all but the simplest classification problems [5, 6]. The object of feature extraction is to isolate the critical features from the T-ray signals to facilitate good classification performance. We note that, after performing the inverse Radon transform, the signals in the time domain can be reconstructed. In this case, a discrete wavelet transform is then performed. Feature extraction then takes place in the wavelet domain, and the final feature vectors have a lower dimension which is more suitable for classification.

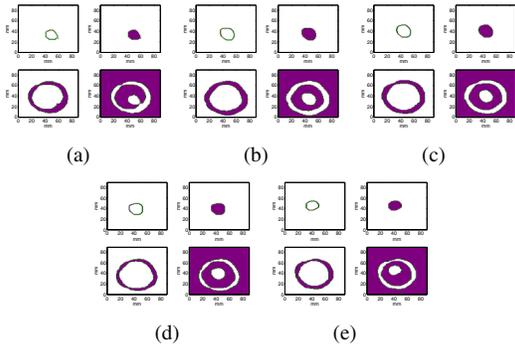


Fig. 5. Segmented image sets at heights of (a) 5 mm; (b) 6 mm; (c) 7 mm; (d) 8 mm and (e) 9 mm from bottom of vial. The four subfigures, clockwise from top left, illustrate the tube edges, tube segments, the air segments (background) and vial segments, respectively.

3.3.1. Discrete Wavelet Transform

Discrete wavelet transform (DWT) is a technique that has a similar importance to FFT in analysing signals. It allows desirable properties of approximation quality concerning analysed signals. Mallat's pyramid algorithm is commonly used in the calculation of discrete wavelet transform coefficients. A pair of mirror low-pass and high-pass filters, are applied iteratively in the computation of wavelet transform.

A step of wavelet transform can be achieved as follows. Project an arbitrary function $f(t)$ to a lower resolution subspace V_j and a detailed subspace W_j according to the following formulae:

$$\begin{cases} c_{j+1,k} = \sum_m h(m-2k)c_{j,m} \\ d_{j+1,k} = \sum_m g(m-2k)c_{j,m} \end{cases} \quad (1)$$

where $f(t)$ represents the signals to be analysed, and $f[m] = c_{0,m}$. Here, $c_{j+1,k}$ are the scaling coefficients corresponding to the space V_{j+1} , while $d_{j+1,k}$ are the wavelet coefficients corresponding to the space W_{j+1} , with j indicating the decomposed wavelet scale [3]. The filters corresponding to the Daubechies-8 wavelet, BiorSpline-6.8 wavelet and Coiflet-3 wavelet are employed in this paper.

Fig. 7 shows the calculation of the DWT using a digital filter bank with an application to feature extraction, where h and g represent the low- and high-pass filters, respectively; \mathbf{c} and \mathbf{d} represent the vectors of coarse and detailed wavelet coefficients.

3.3.2. the Extraction of Wavelet energy coefficients features

In the current approach, the feature extraction process involves several steps. First, the classified pixels are pre-processed by

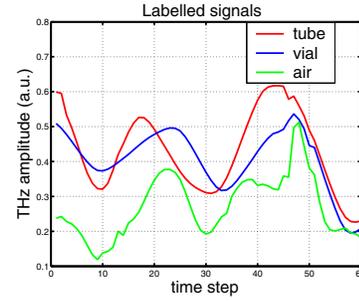


Fig. 6. Illustrate the labelled signals from three classes: the vial, the tube and the back ground, with a height equaling to 7 mm from bottom to top and a time step equaling to 0.15 ps. The signals are truncated in the 60th time step for clarity. The signals for three classes—the vial, the tube and the air (background) have been well separated.

wavelet shrinkage denoising [7] via 'heuristic' SURE algorithm. Then, a two, three or four level discrete wavelet transforms (DWTs) are performed, using the biorthogonal BiorSpline wavelet of order 6.8 and orthogonal Daubechies wavelet of order 8, together with non-orthogonal Coiflet wavelet of order 3, separately, for the target experimental data. Next, Mallat's pyramid algorithms are used to calculate the coefficient vectors \mathbf{c} and \mathbf{d} for labelled signals from each target class, respectively, and therefore wavelet energy coefficients \mathbf{c}^2 and \mathbf{d}^2 are obtained. These calculations are repeated on all classified pixels from each test sample. Finally, the average of the wavelet coarse and detailed energy coefficients are computed for each wavelet subband. The modeling coefficient averages are then joined to produce feature vectors with a dimension of 2, 3 or 4.

The use of wavelet transforms aims to reduce the dimension of T-ray pulses. Owing to the dimension variety of wavelet coefficients resulting from the various wavelet transform levels, averaging the wavelet energy coefficients changes a varying dimension of wavelet energy coefficients to a fixed number of dimensions, which leads to a constant feature vector structure.

3.4. Classification

The current classification experiment is realised by a Mahalanobis distance classifier.

3.4.1. Mahalanobis distance classifier

The quantity d_i in the following equation,

$$d_i(x) = \sqrt{(\mathbf{x} - \mathbf{m}_i)^T C_x^{-1} (\mathbf{x} - \mathbf{m}_i)} \quad (2)$$

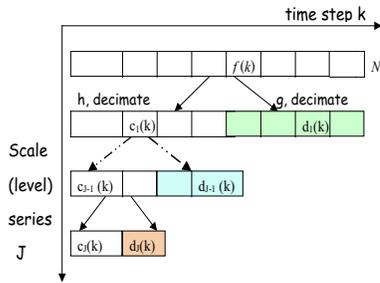


Fig. 7. Fast Wavelet Transform (or Mallat’s algorithm) is the calculation of the DWT using a digital filter bank, which is employed for the feature extraction.

is called the Mahalanobis distance from the feature vector \mathbf{x} to the mean vector \mathbf{m}_x of given classes, where C_x is the covariance matrix for \mathbf{x} used to normalise each given class [8].

Fig. 8 illustrates the basic principle of Mahalanobis distance classifier. The means (templates) for the C classes are labelled by m_1, m_2, \dots, m_C , and the corresponding covariance matrices are indicated by C_1, C_2, \dots, C_C . A feature vector \mathbf{x} is classified by measuring the Mahalanobis distance from \mathbf{x} to each of the means, and \mathbf{x} is assigned to the class for which the Mahalanobis distance is minimum.

3.4.2. The design of classification experiment

The current T-ray CT classification scheme consists of several stages. First to achieve T-ray functional image in the different heights for the vial and the tube set up. Then these images are segmented into different regions, and each pixel location is grouped into a nominal class for five heights. In turn, feature extraction is performed on randomly selected pixels using a wavelet-based algorithm, with the aim of reducing dimensionality to permit effective classification. The features of labelled pixels at one height were used to train the classifier. The pixels at other heights were classified via a Mahalanobis distance classifier and the resultant classification accuracy was measured. The segmentation techniques were achieved via a 2D wavelet scale correlation method as mentioned above.

There are 150 pixels selected from 89×89 image data at each height, at heights from 5 mm to 9 mm, for the calculation of feature vectors, in order to keep the same number of feature for each class. These 150 pixels come from the three different classes, namely the vial, the tube and the air regions. A set of 150 pixels in the T-ray CT image at a height of 7 mm were used for training. Among those selected 150 pixels, each pixel in the CT image from the three classes corresponding to heights equal to 5 mm, 6 mm, 8 mm and 9 mm was

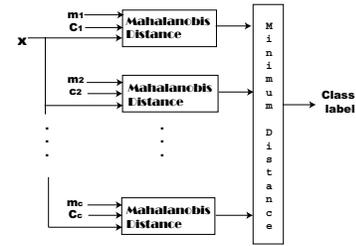


Fig. 8. The basic principle of Mahalanobis distance classifier, (source: www.cs.princeton.edu).

used to test the classifier. The resultant classification accuracy with respect to each test layer was computed via averaging the classification accuracy associated with the 50 tested feature vectors for each class. The number of the time samples which was applied to the DWT was from first 30 to first 350 in steps of 30. In this paper, first 300 time samples are demonstrated for lower calculation expense. A variable number of time samples were chosen to determine the least number samples required for effective classification. The number of time samples used dramatically affects the computation of the algorithm.

4. CLASSIFICATION RESULTS

Table 1 shows the resultant classification performance based on wavelet transforms for feature extraction. The one dimensional (1D) DWT with five wavelet transform scales is investigated. Orthogonal Daubechies 8 (db8) wavelets, non-orthogonal Coiflet 3 (coif3) wavelets and biorthogonal Bior-Splines 6.4 (bior6.4) wavelets are used for the 1D DWT in order to compare their performances. The corresponding classification accuracy ranges, shown in the middle part of each bracket in the table, and the corresponding mean accuracy, shown on the left and right sides of each bracket in the table, at the four test heights and at the combinations of from the first 30 time samples to first 300 time samples are shown in Table 1. In addition, the relative classification accuracy plots are achieved via orthogonal and biorthogonal along with non-orthogonal wavelet bases—db 8 wavelet, bior 6.4 and Coiflet 3 wavelet are illustrated in Fig. 9, which show the expected range of the classification accuracy versus the different number of time samples at the different object heights.

The experimental results show that, for each wavelet family and at each height of the test vectors, the classification accuracies between the first 100 and the first 200 time samples are relatively high, exclude Daubechies 8 wavelets, where the high classification accuracy can be kept till first 250 time samples. This is due to the high signal to noise ration (SNR) of original measured T-ray pulse response in the corresponding

Table 1. This table shows the classification accuracy (%) range, located on the left and right sides of each bracket, and the average classification accuracy, indicated in the middle of each bracket, which are calculated at the four object heights and at the combinations from first 30 time samples to first 300 time samples in time step of 30. The wavelet families involved in this experiment are the db8 wavelet, bior6.8 wavelet, coif3 wavelet.

| | 5 mm | 6 mm | 8 mm | 9 mm |
|---------|----------------|----------------|----------------|----------------|
| db8 | [67, 77.2, 82] | [91, 93.2, 95] | [79, 86.1, 93] | [74, 80.2, 86] |
| bior6.8 | [72, 78.2, 87] | [90, 93.3, 97] | [83, 87.3, 90] | [77, 81.6, 90] |
| coif3 | [71, 78.2, 87] | [85, 93.3, 99] | [80, 87.3, 94] | [77, 81.6, 88] |

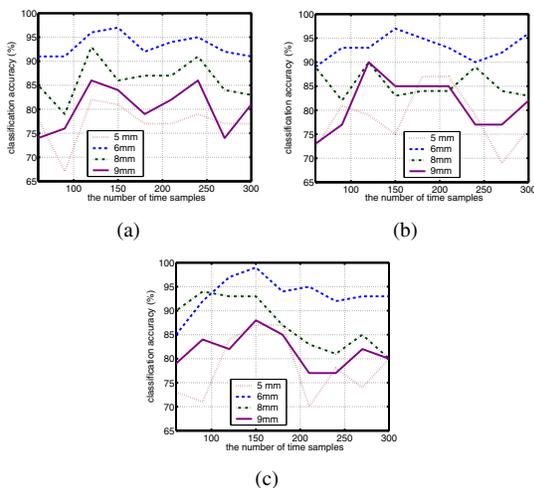


Fig. 9. Plot the classification accuracies at four object heights equal to 5 mm, 6 mm, 8 mm and 9 mm versus the different number of time samples. The classification is realised via (a) db8 wavelets, (b) bior6.8 wavelets, (c) coif3 wavelets.

time range compared to other time range, please see Fig. 2. The classification accuracy curves demonstrated in Fig. 9 can be approximately thought of being changeable around average classification accuracies, shown in the middle part of each bracket in Table 1. The average values of the classification accuracies corresponding to each wavelet family are similar, which are decreased with object heights from 6 mm, 8 mm, 9 mm and 5 mm. The far more distance between the heights the train vectors and test vectors are measured, the lower average classification accuracies are calculated, vice versa. Fig. 9(a) displays the obvious variety of classification accuracy curves at four test heights for Daubechies 8 wavelets. Fig. 9(b) exhibits the classification accuracy curves at four test heights for bior 6.8 wavelets, where the overlap of classification accuracy happens at the heights of 8 mm and 9 mm, between 100 and 200 time samples and the classification accuracy at the height equal to 5mm is bigger than the classification accuracy at the heights of 8 mm and 9 mm. Fig. 9(c) indicates that a little bit

overlapping of classification accuracy happens at the heights of 5 mm and 9 mm, between 100 and 200 time samples for Coiflet 3 wavelets.

The Daubechies 8 wavelet gives better classification performance of the three types of wavelet functions owing to its obvious change of classification accuracy with the variety of the heights, which correctly reflects the information variety of measured data with height change.

5. CONCLUSION

Extracted features based on wavelet energy coefficients allow expected classification performance, with reduced computation expense, compared to FFT-based techniques, which usually rely on other computationally expensive algorithms, such as Genetic Algorithms, to achieve optimum performance. Wavelet transforms for feature extraction facilitate the real time classification of homogeneous materials in 3D space. The significance of this work is that we can now identify materials embedded within a 3D heterogeneous structure.

6. REFERENCES

- [1] B. Ferguson, S. Wang, D. Gray, D. Abbott, and X.C. Zhang, "Toward functional 3D T-ray imaging," *Physics in Medicine and Biology (IOP)*, vol. 47, pp. 3735–3742, 2002.
- [2] Xiaoxia Yin, Brian W.-H. Ng, Brad Ferguson, Samuel Peter Mickan, and Derek Abbott, "Information fusion and wavelet based segment detection with applications to the identification of 3d target t-ray ct imaging," *IEEE Sensor Letter*, 2006, submitted.
- [3] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, CA., 1999.
- [4] R. C. Gonzalez and R. E. Woods, *Digital Image processing*, Prentice-Hall, Inc., New Jersey, 2002.
- [5] R.J. Schalkoff, *Pattern Classification: Statistical, Structural and Neural Approaches*, John Wiley and Sons, Inc., New York, 1992.
- [6] S. Pittner and S. V. Kamarthi, "Feature extraction from wavelet coefficients for pattern recognition tasks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 1, pp. 1085–1094, 1999.
- [7] D. L. Donoho, "De-noising by soft thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.
- [8] J. Schürmann, *Pattern Classification: A Unified View of Statistical and Neural Approaches*, John Wiley and Sons, Inc., New York, 1996.