A New Method for Processing Passive Sonar Data

Trevor Wood
OCIAM
University of Oxford, UK
woodtm@maths.ox.ac.uk

David Allwright
OCIAM
University of Oxford, UK
allwrigh@maths.ox.ac.uk

Philip Bond
OCIAM
University of Oxford, UK
bond@maths.ox.ac.uk

Stephen Long
Thales Underwater Systems
Cheadle, UK
stephen.long@uk.thalesgroup.com

Irene Moroz
OCIAM
University of Oxford, UK
moroz@maths.ox.ac.uk

Abstract – *Processing passive sonar data involves detecting and monitoring tones present in an audio signal. A new method for processing passive sonar data is proposed. The method includes a pre-processing step using Radon and wavelet transforms, which is shown to improve the receiver operating characteristics of the data. After pre-processing, the extraction of continuously present frequencies from the data is treated as a multitarget tracking problem in the frequency-frequency rate domain. Frequency tracks are extracted using the probability hypothesis density filter from the theory of finite set statistics. Results from application of the method to both simulated and real passive sonar data are presented.*

Keywords: Passive sonar, tracking, wavelets, finite set statistics, PHD filter, radon transform.

1 Introduction

Passive sonar involves listening to, without transmitting, acoustic signals in order to detect other vessels. We consider the case where directional information about the signal is not available so localisation of detected vessels is not possible. The goal is to identify, in the signal, the audio spectrum of any vessels present. This audio spectrum is comprised of important frequencies (or tones) which correspond to, for example, rotating machinery in the engine of the detected vessel. It is hoped that automatic detection and monitoring of the tones present in the signal will ease the workload of the operator and could potentially aid target classification.

This paper is an exposition of a methodology for the processing of passive sonar data. The novelty of the method is two-fold.

Firstly, we introduce a pre-processing step to denoise the data to enable easier detection of tones. The relevant techniques are the Radon transform and wavelet denoising.

The second step is the extraction and monitoring of tones present in the signal. The number of tones present together with their frequencies, which may not be constant, must be jointly determined from the set of potential detections returned from the data. The pre-processed data may include some combination of: unpredictable tone shifts, appearance and disappearance of tones, missed detections of tones, false detections, imprecision in the frequency values detected for tones and crossing tones. Due to these factors, extraction and monitoring is non-trivial and is formulated here as a multitarget tracking problem in the frequency-frequency rate domain.

There are a wide range of available multitarget tracking methods (see, for example, [1]) but we focus on those from the theory of finite set statistics (FISST), largely due to Mahler [3]. FISST methods are chosen because they enable a unified multitarget tracking process, with no need for separate heuristics for data association or track management. In particular, we use the well-known Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter [7], an approximation to the multitarget Bayes filter from FISST.

The pre-processing methodology is presented in Section 2 and the multitarget tracking is discussed in Section 3. The methodology is tested on both real and simulated data in Section 4 and conclusions are presented in Section 5.

2 Pre-Processing

The aim of the pre-processing step is to denoise the data to allow accurate extraction of potential detected tones. This process is aided by two nonlinear transformations, the Radon and wavelet transforms, motivated by the idea that separation of signal from noise will be easier in the transform domain. The denoising process consists of transforming the data, denoising the data in the transform domain and then inverting the transform.

2.1 Radon and Wavelet Transforms

i) Radon Transform - Let f be a continuous function vanishing outside of some disc in $\mathbb{R}^2$. The Radon transform, denoted $\mathrm{R}f$ and defined on the set of lines in $\mathbb{R}^2$, is given by:

$$Rf(L) = \int_L f(x) d\sigma(x) \tag{1}$$

where the integration is performed with respect to the arclength measure $d\sigma(x)$ on L. Thus, the Radon transform has large values for lines which are strong in the original data. Applying a Radon transform to the data exploits the fact that we expect the tones generally to be straight lines in frequency-time space. As a result, the signal energy will be more concentrated in the Radon transform domain whereas noise should not be, making separation of signal and noise easier.

It should be noted that previous work by Sun and Willett [6] considered use of the Hough transform, which is closely related to the Radon transform, for detecting tones in passive data. The key difference is that the Radon transform permits use of amplitude information whereas the Hough transform is discrete and forces thresholding first.

ii) Wavelet Transform - A wavelet basis is chosen for the data, with the basis divided into levels representing different scales of detail. Elements of the wavelet basis which correspond to smaller scale detail are then subject to thresholding. This is analogous to expressing a function in terms of its Fourier series and then thresholding the higher frequency coefficients to remove high frequency noise. The wavelet basis used is 'sym4'. The method for choosing threshold levels is given in [2]. 'A Wavelet Tour of Signal Processing' [4] is a good general purpose reference book for this subject.

2.2 Examples for Simple Simulated Data

Wavelets are a popular method for image processing. We present two examples; the first demonstrates the advantage of employing the Radon transform and wavelet denoising (Radon+wavelets) as compared to only using wavelet denoising. The second example shows that the Radon transform together with a less sophisticated thresholding method will run into problems.

For the first example, Fig. 1 shows two constant tones in the presence of additive white Gaussian noise. This data is then denoised in two ways; firstly, using the Radon transform followed by wavelet denoising (and finally an inverse Radon transform) and secondly, using only wavelet denoising.

Fig. 2 shows the Radon transform of the raw data. The two lines visible in the raw data are now represented by two points in the radon transform domain with angle of projection 0/180 degrees due to the vertical slopes of the lines in the frequency time domain.

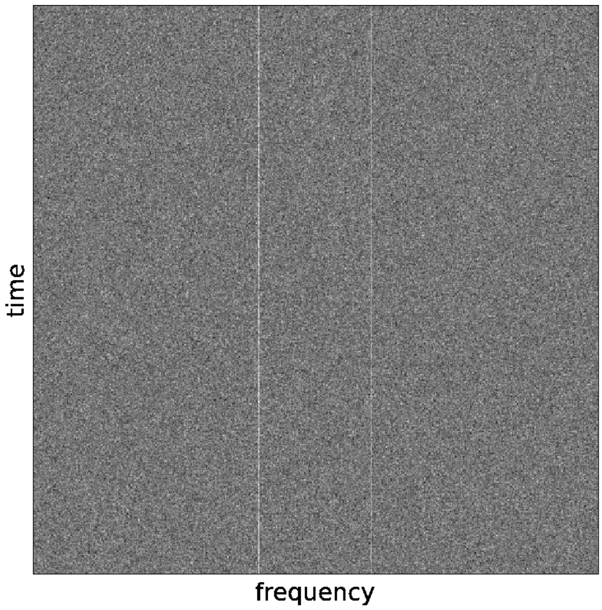

Figure 1: Raw data with two constant important frequencies plus noise

Viewing the data in both the standard and Radon domains (as in Fig. 1 and Fig. 2 respectively), it becomes clear that separating signal from noise might be easier in the latter domain.

In this example the signal to noise ratio (SNR) in the decibel scale over the whole original data was 2.8dB. After processing using only wavelets this had increased to 4.4dB, but after processing with Radon+wavelets there was a more substantial increase to 15.4dB. This simple example demonstrates the benefit of using both Radon and wavelet transforms.

A more complicated example is shown in Fig. 3. Here, there are four tones; two of them shifting, one constant and one unstable (this is the wavering tone third from the left). In this example, we compare the results of Radon+wavelets against a method using the Radon transform with naive thresholding, rather than with wavelet thresholding, in order to ascertain if wavelets are necessary at all.

Fig. 4 shows the output of Radon with naive thresholding. This method is not able to pick up some wrinkles of important detail, such as the briefly present tone on the far right. Furthermore, if a line is present for part of the image, the Radon with naive thresholding tends to declare this line for the whole image, resulting in artifacts. By contrast, Fig. 5 shows the output of Radon+wavelets in which there is a perceptible improvement in the image with respect to the contrast of the tones against the noise. From these preliminary results, we expect the best pre-processing to be due to the use of both Radon and wavelet transforms. This claim will be tested on less simplistic data in Section 4.

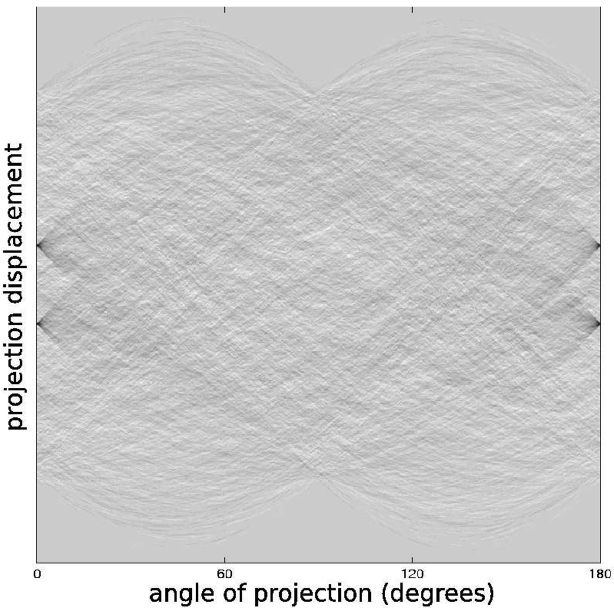

Figure 2: Radon transform of data shown in Fig. 1.

3 Tracking

After pre-processing the data, the aim is to establish, at each time step, how many tones are present in the data and the frequency of each tone. Associations between frequencies at different time steps are also required to enable, for example, detection of Doppler shifts.

Potential detections for tones are extracted from the processed data. Despite the denoising step, there are still likely to be some imperfections.

Some features of the problem include:

- Uncertain dynamics of true tones (as frequencies may unexpectedly shift due to Doppler or engine effects)
- Possible appearance or disappearance of tones as vessels come into or out of range
- Missed detections of true tones due to imperfection of data or other factors, such as the Lloyd's mirror effect
- False detections (possibly many more of these than detections of true tones)
- Imprecise frequency measurements for the detected tones (due to imprecision in equipment, diffuse tones, problems with Fourier transforms on noisy signal)

The problem is formulated as a multitarget tracking problem using the framework of FISST in the following way. Let X_k be the set of tones present at time step k,

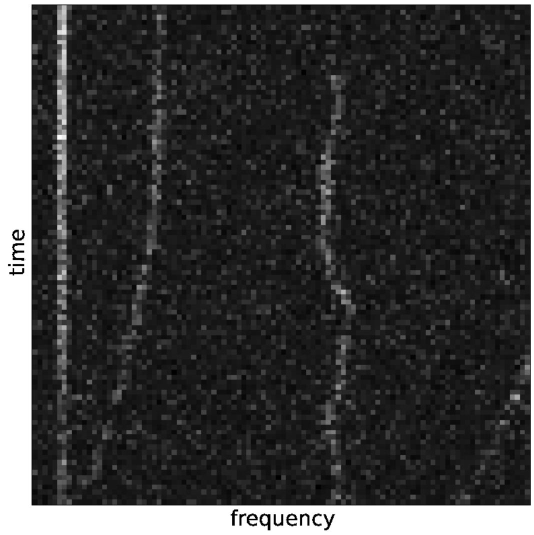

Figure 3: Raw data with four important frequencies plus noise

Z_k be the set of detections at that time step. Thus,

$$X_k = \{\mathbf{x}_k^{(1)}, ..., \mathbf{x}_k^{M_k}\} \tag{2}$$

$$Z_k = \{z_k^{(1)}, ..., z_k^{N_k}\} \tag{3}$$

$$\mathbf{x}_k^{(j)} = (f_k^{(j)}, \dot{f}_k^{(j)})^T \tag{4}$$

where $f_k^{(j)}$ and $\dot{f}_k^{(j)}$ are frequency and frequency rate of tone $\mathbf{x}_k^{(j)}$, M_k and N_k are the number of tones and detections at time k respectively.

X_k and Z_k are random finite sets, which means that their members are random and the number of members in each set is also random. This represents the fact that the number of tones is unknown and their frequencies are also unknown. We use the model

$$X_k = \left(\bigcup_{\mathbf{x} \in X_{k-1}} S_{k|k-1}(x) \right) \cup \Gamma_k \tag{5}$$

$$Z_k = K_k \cup \left(\bigcup_{x \in X_{k-1}} \Theta_k(\mathbf{x}) \right) \tag{6}$$

where, $S_{k|k-1}$ represents tones surviving from time step $k-1$, and Γ_k represents new tones appearing. The model for surviving tones is:

$$S_{k|k-1}(\mathbf{x}) = \begin{bmatrix} 1 & dt \\ 0 & 1 \end{bmatrix} \mathbf{x} + \mathbf{v}_k \tag{7}$$

where $\mathbf{v}_k$ is noise representing the uncertainty of future changes in the tone. This noise is usually assumed to be zero-mean Gaussian. dt is the time step.

K_k represents false detections and $\Theta_k(\mathbf{x})$ represents detection from true tones and is given by

$$\Theta_{k|k-1}(\mathbf{x}) = \begin{cases} \mathbf{x} \cdot (1,0) + w_k & \text{with prob } p_D \\ \emptyset & \text{with prob } 1 - p_D \end{cases} \tag{8}$$

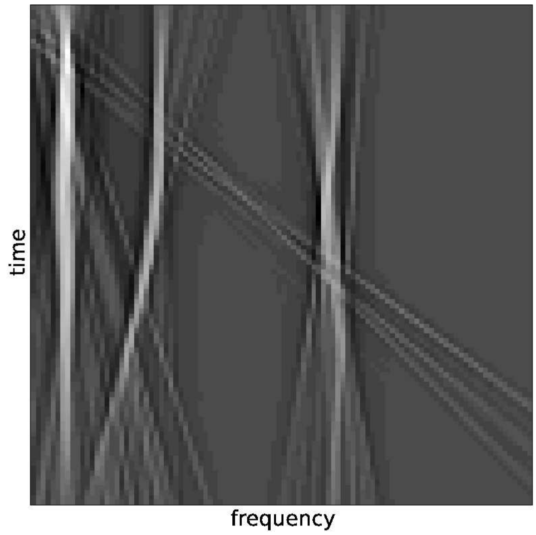

Figure 4: Data from Fig. 3 after Radon transform and naive thresholding

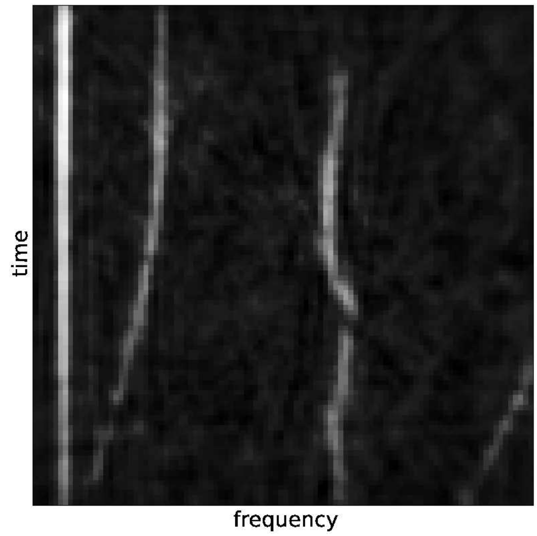

Figure 5: Data from Fig. 3 after Radon and wavelet denoising

where w_k is noise (also assumed to be zero-mean Gaussian) representing the imprecision of the frequency in the detections and p_D is the probability of detecting the tone, which may be estimated as described in Section 4.1.

The multitarget Bayes filter of FISST allows the probability density for X_k, conditioned on Z_k, to be computed sequentially for all k. This enables a unified multitarget tracking process, with no need for separate heuristics for data association or track management. More details are available in [3].

The multitarget Bayes filter from FISST is computationally intractable and hence must be approximated. Propagating the 1st moment of the multitarget posterior, called the probability hypothesis density (PHD), gives the PHD filter. The PHD filter itself must also be approximated for tractability, and two known techniques use sequential Monte Carlo methods (SMC-PHD) and Gaussian mixtures (GM-PHD), as described in Vo et. al 2005 [8] and Vo & Ma 2006 [7] respectively.

The GM-PHD filter has favourable computational characteristics, such as linear complexity with respect to the number of detections and targets. Furthermore, it is demonstrated in [9] that its performance is comparable to that of the Multiple Hypothesis Tracker (MHT), widely considered the state of the art [1]. The GM-PHD filter implementation used will be in the form proposed by Panta et. al [5], who use a track labelling system to allow association of tracks between time steps.

Using tracking methods enables incorporation of available information about the data in the form of explicit statistical models for all relevant phenomena. Tracking methods are designed to be robust to noise in the data and the set of tracks returned is a concise expression of the important information contained in the original data.

As an example, Fig. 6 shows the output of the GM-PHD filter for the processed data from Fig. 3.

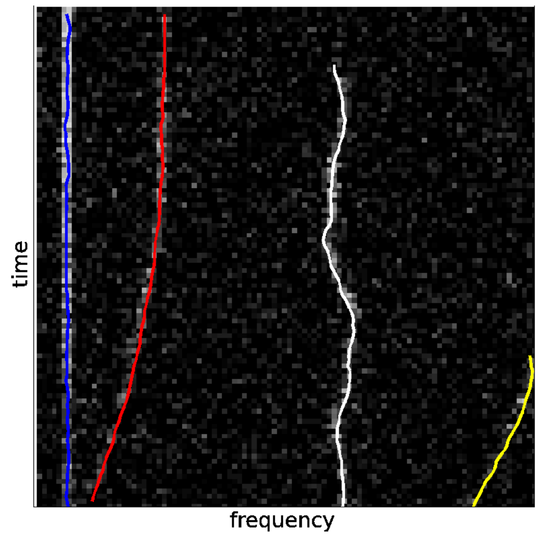

Figure 6: Frequency tracks extracted from the data shown in Fig. 3 using the GM-PHD filter.

3.1 Extra Complications for Passive Data

There are some extra features of passive data that are somewhat different to a standard tracking scenario that will need to be addressed. Firstly, it is possible for the signal to fade intermittently for some tones due, for example, to the Lloyd's mirror effect, whereby sounds arriving via different paths may interfere constructively or destructively, causing amplified signal or loss of signal respectively. In the latter case, detection of the affected tones may be impossible for dozens of time steps. In a normal tracking scenario, the lack of detection of a target for a large number of consecutive time steps in a row would be cause for concluding that the target had disappeared. To prevent loss of tracks on such tones, we propose a simple track linking method.

If a track on a tone is deleted due to missed detections at time step n_1, at which point it had frequency f_1 and frequency rate $\dot{f}_1$, and another tone appears at time step n_2 with frequency f_2, the tracks are linked if $|(f_1 + (n_2 - n_1)\dot{f}_1) - f_2| < T$ and $0 < n_2 - n_1 < U$, where T and U are adjustable parameters.

A second difference is that tones can cross over in the 1-D frequency domain more easily than targets in 3-D domains. This could lead to confusion about which tones before the crossing correspond to which tones after the crossing. To alleviate this problem, a simple solution would be to identify intersections and associate the tones with closest frequency rates before and after.

4 Results

The new method was tested on two sets of data. The first is a test scenario designed for the performance assessment of existing algorithms made available thanks to Thales Underwater Systems. The scenario includes many aspects which might be present in real passive sonar data including:

- Doppler and non-Doppler related frequency shifts
- Crossing tones
- Intermittently fading contacts
- Unstable tones
- Diffuse tones

The scenario is based on a target trajectory hence the Doppler shifts will be representative of an actual target motion. The crossing of tones is achieved by ensuring that a tone with a non-Doppler related shift crosses a tone which is Doppler related. For example, speed changes will cause an engine related tone to undergo significant changes in frequency. Such frequency changes could cause some of the tones from the same vessel, but from a different source within that vessel, to cross.

The raw data is shown in Fig. 7 and the data after pre-processing is shown in Fig. 8 for comparison. Tones that are visible but varying in strength in the raw data become more consistent after pre-processing, while other tones that were difficult to detect by sight become more pronounced. It can also be seen that noise level has generally been reduced though the Lloyd's mirror effect in the top left corner of the image is still visible.

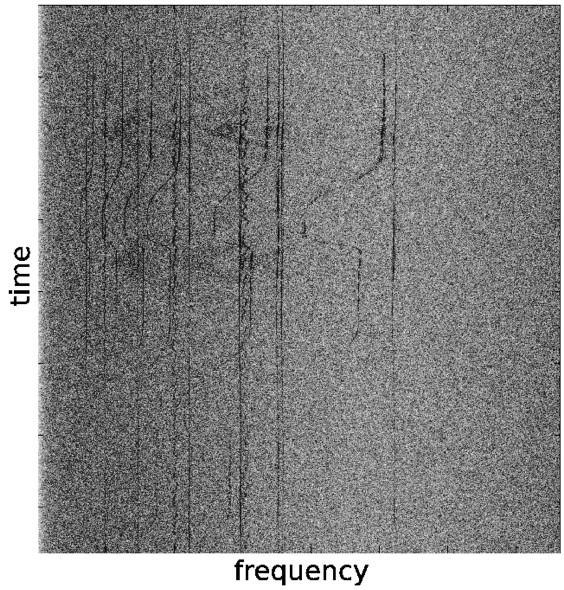

Figure 7: Simulated Test Scenario

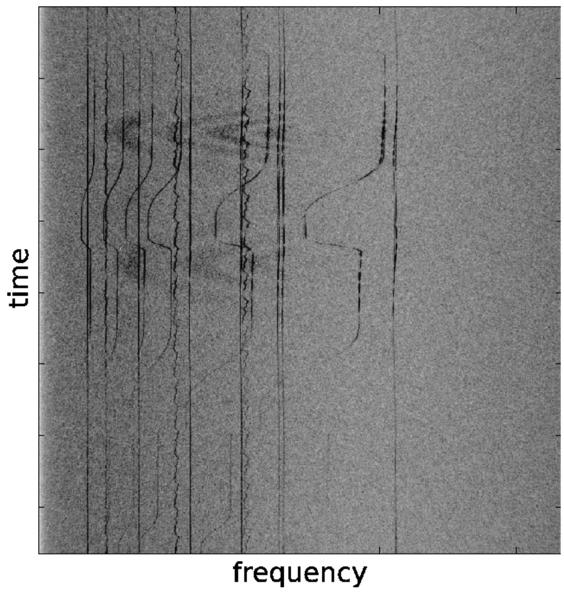

Figure 8: Simulated Test Scenario after pre-processing

The second set of data is taken for a real passive sonar recording of a Diesel engine. The raw data is shown in Fig. 9 and the pre-processed data in Fig. 10. In contrast with the first example, the noise in this data is less uniform making this a more realistic test of the method. Fig. 10 shows that the strong noise on the right hand side of the image is not eliminated by pre-processing but that many tones not visible in the raw data become apparent despite the remaining noise.

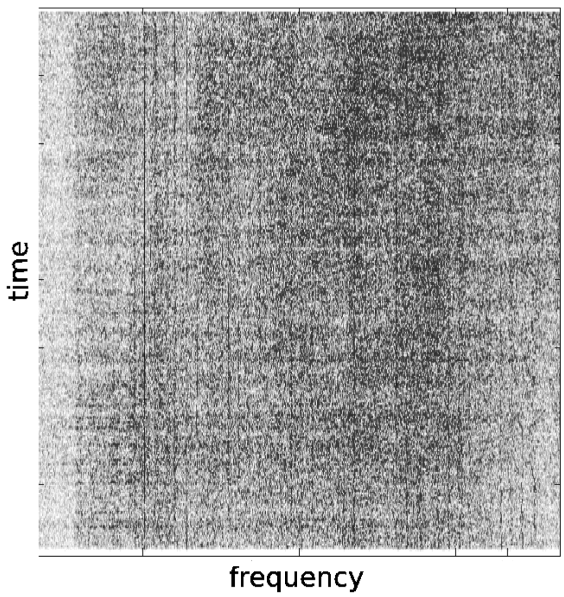

Figure 9: Diesel engine passive data

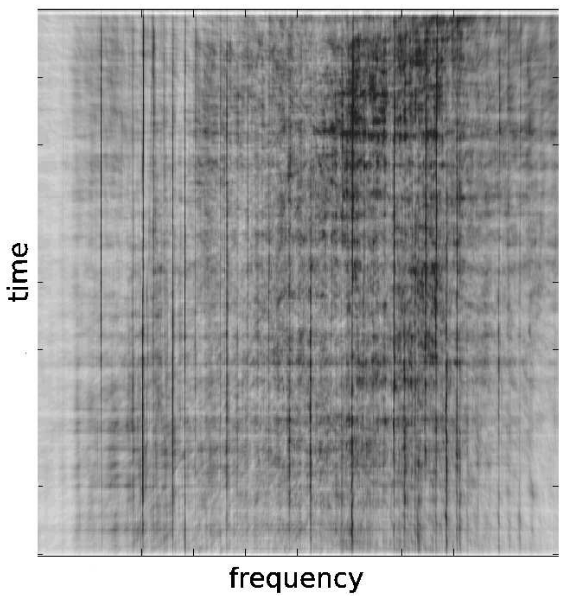

Figure 10: Diesel engine data after pre-processing

4.1 Effect of Pre-Processing Step

As discussed in Section 3, the process of extracting detections of tones from the data may be expected to lead to false detections as well as missed detections of real tones. We monitor the levels of both false detections and missed detections as a test of the effectiveness of the pre-processing techniques. For the test, a subset of the data from Fig. 7 was chosen, and the true tones present were identified. Detections were extracted using various thresholds and the proportion of true tones correctly reported was recorded along with the level of false detections for each threshold. This test was repeated for three cases: i) raw data, ii) wavelet thresholding and iii) Radon+wavelet thresholding. The results were used to measure the trade off between the probability of missing true tones and the probability of detecting false ones in each case. This trade off between probability of detection and probability of false alarm is closely linked to the idea of receiver operating characteristics (ROC).

Fig. 11 shows this trade-off and is analogous to a ROC curve. It shows that a significant improvement in the level of false detections for any probability of detection required can be achieved by using Radon+wavelet pre-processing. The results also confirm the superiority of Radon+wavelets as opposed to wavelets alone.

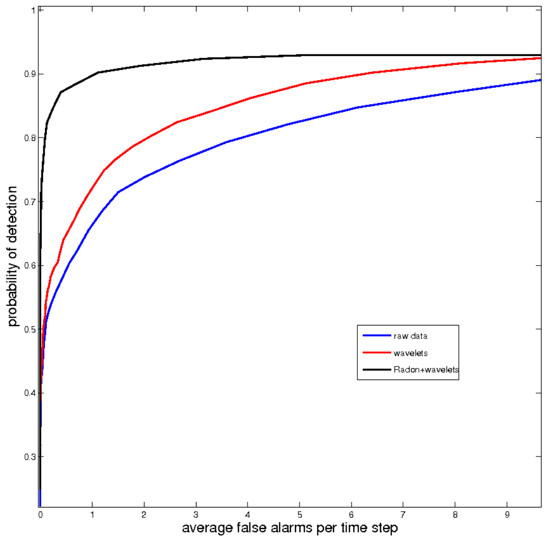

Figure 11: Average level of false alarms vs probability of detection for various pre-processing methods for a subset of the data from Fig. 7

4.2 Tracking step

Fig. 12 and Fig. 13 show the output of the tracking process for the simulated and real data respectively.

For the simulated data, all 16 tones known to be present are detected and largely maintained despite extended sections of fading and crossovers. However, the output is not perfect and there are some sections where fading combined with crossovers has resulted in mistaken track association.

For the real data, 21 tones are identified and tracks maintained successfully on all of them. As the data is real, the 'ground truth' is not known but the tracks reported all correspond to peaks in the FFT of the raw data (a potential check only in this special case in which the frequencies appear constant for the whole time period). Furthermore, a close inspection of the

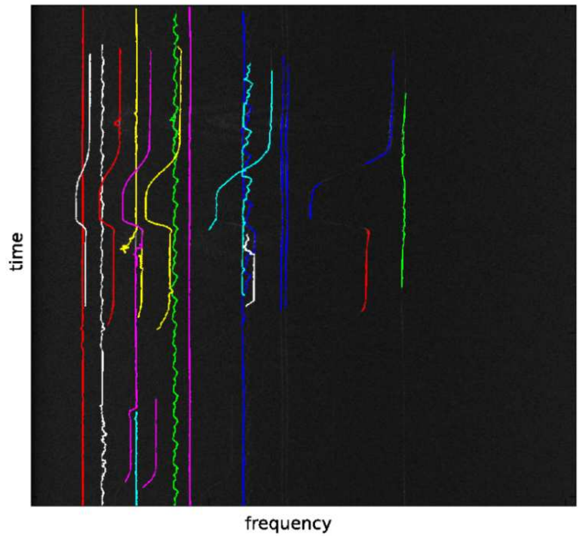

Figure 12: Tracker output for the simulated data

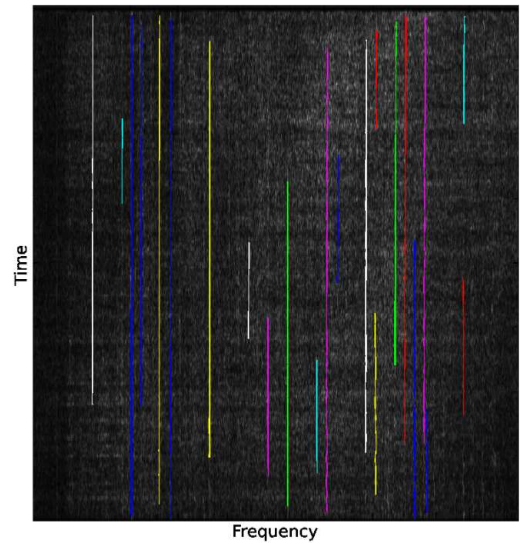

Figure 13: Tracker output for the diesel data

data confirms the tones and the periods for which they are declared are plausible.

5 Conclusions

A new method for processing passive sonar data has been presented. An initial implementation on both real data and a simulated complex scenario shows promising results. The Radon and wavelet transforms as well as the tracking techniques from the theory of finite set statistics are potentially valuable tools for the processing of passive sonar data. Future work comparing the new method to existing methods for frequency detection in passive sonar data would be valuable. A rigorous linking between passive and active sonar with both using finite set statistical methods would also be of interest.

Use of the Radon transform exploits the fact that tones tend to move in straight lines in frequency-time space. Future implementations could improve performance by using more general information about the behaviour of tones, incorporated by using a more generalised form of the Radon transform.

Using this method, phenomena, such as Doppler shifts, can be automatically detected. A natural extension to this work would be to to link this information back to the tracking step, so that detection of Doppler shifts can be used to inform beliefs about the target's dynamics and vice versa.

Acknowledgements

TW's work is generously funded by an EPSRC CASE award with Thales Underwater Systems and Thales Aerospace.

References

[1] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, 1999.

[2] D.L. Donoho "De-noising by soft-thresholding" *IEEE Transactions on Information Theory*, Vol 41, pp. 613-627, 1995

[3] R.P.S. Mahler, *Statistical Multisource-Multitarget Information Fusion*, Artech House, 2007.

[4] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 1999.

[5] K. Panta, B. Vo, and D.E Clark "An efficient track management scheme for the Gaussian-mixture probability hypothesis density tracker," *Fourth International Conference on Intelligent Sensing and Information Processing*, pp. 230-235, 2006

[6] Y. Sun and P. Willett "Hough transform for long chirp detection", *IEEE Transactions on Aerospace and Electronic Systems*, Vol 38, pp. 553-569, 2002.

[7] B. Vo, and W. Ma "The Gaussian mixture probability hypothesis density," *IEEE Transactions on Aerospace and Electronic Systems*, Vol 54, pp. 40914104, 2006.

[8] B. Vo, S. Singh and A. Doucet "Sequential Monte Carlo methods for multi-target filtering with random finite sets," *IEEE Transaction on Aerospace and Electronic Systems*, Vol 41, pp. 1224-1245, 2005.

[9] T.M. Wood, "Tracking in dense clutter with the phd filter," *Proceedings of the IMA Mathematics in Defence Conference*, Farnborough, 2009