Dolphin Whistle Frequency Estimation using Gaussian Mixture Probability Hypothesis Density Filter

Imtiaz Ahmed*

Department of Applied Physics, Electronics & Communication Engineering, Faculty of Engineering and Technology, University of Dhaka, Dhaka-1000, Bangladesh

Abstract

This paper formulates the automation of dolphin whistle track estimation process as a Multiple Target Tracking (MTT) problem using Random Finite Set (RFS) approach. It focuses on achieving possible automation in dolphin whistle tracking using the Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter. Acoustic recordings of three different dolphin species have been considered. Simulation results corroborate that automation in dolphin whistle tracking has been achieved. The GM-PHD filter has been able to produce reliable estimate of whistle frequencies in the presence of multiple whistles, spontaneous death/birth of whistles and multiple whistles crossing each other.

Keywords: Multi-target Tracking (MTT), Random Finite Set (RFS), Probability Hypothesis Density (PHD), Gaussian Mixture Probability Hypothesis Density (GM-PHD) Filter, Dolphin Whistle Tracking

IJCVSP
International Journal of Computer Vision and Signal Processing

ISSN: 2186-1390
http://www.ijcvsp.com
Article History:
Received: 7 August 2012
Revised: 1 October 2012
Accepted: 21 October 2012
Published Online: 25 October 2012

1. INTRODUCTION

Passive Acoustics Monitoring (PAM) technique is widely used in detection and species classification of dolphins. This allows advancement of general knowledge of dolphin species identification and supports conservation and mitigation efforts [1]. Acoustic techniques are complementary to visual detection and classification of dolphin species [2]. Dolphins have a range of vocalization capabilities and they use them for different purposes [3]. Their impressive vocalization capabilities can be categorized into three classes (i) broadband short duration clicks, (ii) broadband pulsed sounds (iii) continuous narrow band frequency modulated whistles. Dolphin whistles have been shown to contain species specific information [4]. This fact led to the use of whistle tracks extracted from hydrophone measurements in dolphin species classification algorithms. Most of the methods of automation in whistle track are mainly based on spectrogram techniques [5]. A completely different approach can be proposed based on the Random Finite Set (RFS) formulation. The RFS [6] approach is an emerging and promising technique in the field of Multiple Target Tracking (MTT). Mahler proposed a natural, elegant and rigorous Bayesian framework [7] based on the RFS for the theory of MT filtering. Since then a number of efficient and computationally tractable approximations to MT Bayes filter have been developed.

One such filter is called the Probability Hypothesis Density (PHD) filter [8], [9], [10] which jointly estimates the state and the number of targets from a set of noisy measurements. But the recursion in PHD filter requires solving multi-dimensional integrals that do not, in general, have closed-form solutions [11]. When the closed form solution does exist, the Gaussian-Mixture PHD (GM-PHD) filter provides it [12].

The purpose of this article is to automate the whistle track extraction process using the GM-PHD filter.

2. DIFFICULTIES IN DOLPHIN WHISTLE FREQUENCY TRACKING

Performance of the species identification algorithms relies on the accuracy of the whistle track estimation process. The challenges in tracking the frequency contents of dolphin whistle can be attributed to: (i) death/birth of frequency tracks which means that number of whistles

*Corresponding author
Email address: imtiaz@univdhaka.edu (Imtiaz Ahmed)

is unknown and generally varies over time (ii) presence of large amount of background noise in the recorded signal (iii) presence of multiple whistles cross each other. Based on challenges mentioned in (i) and (iii), estimation and tracking dolphin whistle can be regarded as MTT problem. In this work the potential of the RFS approach to tackle these challenges has been investigated.

3. A SURVEY ON AUTOMATION IN WHISTLE TRACK DETECTION METHODS

A semi-automated whistle contour extraction method has been proposed which works if the start and end points are known as a-priori and the SNR of acoustic recording is very high [13]. A noise removal method to facilitate contour extraction in natural waters has been proposed [14] where dolphin echolocation clicks are removed by the sequential application of a vertical edge suppression filter and an exponential smoothing filter. The start and end points of whistles are identified by drops in local SNR. A different approach [15] to denoising has been proposed which combines the outputs of four different two-dimensional filters applied to the signal spectrogram in order to reduce the noise level. This method is based on the assumption that dolphin whistles are smooth curves without sudden jumps in frequency. This method can only extract a single whistle at each time, so it is not appropriate for tracking multiple whistles. Another approach to automate detection and frequency estimation of dolphin whistles can be found in [16]. The method presented there is based on Adaptive Notch Filters (ANFs) and has been applied to bottlenose dolphin whistle tracking. Another fully automated system for whistle tracking, Real-time Odontocete Call Classification Algorithm (ROCCA), has been developed to allow real-time acoustic species identification [2]. ROCCA automatically extracts the whistle contour from the wav file by stepping through the file one FFT window at a time. The fundamental frequency of the whistle contour is selected based on the peak frequency in each window.
RFS can be used to formulate multiple dolphin whistle track estimation. The potential of the RFS formulation in automation of whistle tracking from the raw hydrophone measurements has been investigated here.

4. THE GM-PHD FILTER IN WHISTLE TRACKING

4.1. THE PROBABILITY HYPOTHESIS DENSITY

The PHD function $\nu(\cdot)$ is the first order moment of the target posterior density function. In the context of MTT, the position and the number local maxima of the posterior density function represent the probable target position and number. Hence PHD function can be used to generate estimates of the target positions and number [17].

4.2. THE GM-PHD FILTER RECURSION

The GM-PHD filter approximates the target posterior density as a Gaussian Mixture Density function. It propagates an intensity using the PHD recursion. The recursion consists of prediction and update steps and forms the basis of a general MTT algorithm, called the GM-PHD filter [12]. During prediction step, the GM-PHD filter uses the system model to predict mean and covariance of the Gaussian Mixture. Predicted mean and covariance are updated using the likelihood function once the measurements are available. This update step requires a measurement model specifying the likelihood function. The system and measurement model define the state space model for the filtering problem under consideration. The recursion in GM-PHD filter causes the number of Gaussian components resulting from the recursion step to increase without bound. Hence at the end of each recursion pruning and merging operations is performed. Pruning removes Gaussian terms of low weights and keeps a certain number of terms of the strongest weights. If the distance between the Gaussian components are below a certain threshold, they are merged into a single Gaussian density.

4.3. DYNAMICAL MODEL FOR DOLPHIN WHISTLE TRACKING

In MTT context, different dolphin whistles represent different targets. The state of the MTT in this essence consists of the values of frequencies along with other parameters like chirp rate and possibly higher derivatives of frequency. In this work state x_k at time step k consists of frequency f and chirp rate α:

$$x_k = [f\ \alpha]^T \tag{1}$$

The chirp rate α defines the time rate of change of whistle frequency and has the unit Hz/s. A linear Gaussian discrete state space model is used for dolphin whistle tracking:

$$x_k = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} x_{k-1} + v_k \tag{2}$$

where v_k is a zero mean Gaussian noise with a diagonal covariance matrix, so that system noise of the frequency and sweep rate are uncorrelated with different variances i.e. Σ_v, f_s is the sampling frequency related with sampling period $T = \frac{1}{f_s}$ .The state transition density $f_{k|k-1}(x_k|x_{k-1})$ is $\mathcal{N}(x_k; m_{f_{k|k-1}}, \Sigma_{f_{k|k-1}})$ where $m_{f_{k|k-1}} = 0_{1\times 2}$ and
$\Sigma_{f_{k|k-1}} = \begin{bmatrix} \sigma_f^2 & 0 \\ 0 & \sigma_\alpha^2 \end{bmatrix}$.

Only frequency measurements are available, and the measurement model is:

$$z_k = Hx_k + w_k \tag{3}$$

where the measurement matrix H is $H = [1\,0]$ and w_k is zero mean Gaussian noise with covariance Σ_w. Hence the likelihood function $\mathcal{L}(z_k|x_k)$ is a Gaussian density $\mathcal{N}(x; m_{\mathcal{L}}, \Sigma_{\mathcal{L}})$

where $m_{\mathcal{L}} = 0_{1\times 1}$and $\Sigma_{\mathcal{L}} = \Sigma_w$. The reason behind choosing this model is that such model has been used for tracking a single whistle track using Particle filter approach in [18] and results obtained were impressive. A different model involving next higher order derivative to chirp rate in the state matrix, has been used in [15]. Such approach used the Kalman filter (KF) to track the dolphin whistle. But the RFS approach is a much more refined approach than the classical KF and has the potential to produce accurate state estimates using only the frequency and the chirp rate.

4.4. MODEL FOR SPONTANEOUS BIRTH OF WHISTLE FREQUENCY

The spontaneous birth of whistle frequency components can appear in the frequency range of 2 kHz– 30 kHz [2]. The probability of birth of the track is high in the vicinity of the measurements [19], [20]. Hence the birth of tracks is modelled with a Poisson RFS whose intensity γ_k is a Gaussian mixture density:

$$\gamma_k = \frac{1}{|Z_k|} \sum_{z \in Z_k} \mathcal{N}(x; [z, 0], \Sigma_v) \tag{4}$$

The weights of the individual Gaussian components in this mixture density is chosen to be $\frac{1}{|Z_k|}$. The spreading of each component in the mixture is controlled by the covariance Σ_v of state transition density.

4.5. MEASUREMENT GENERATION PROCESS

Implementation of the GM-PHD filter requires a set of measurements at discrete instants of time for joint estimation of the state and the target number. These measurements are generated from raw audio data from hydrophone using whistle contour detection based on a modified spectrogram [21]. The aims of the modifications are two-fold. First aim is to reduce the effect of echolocation clicks which are commonly present in recordings of dolphin whistles. The second process is a normalization to equalize the signal with respect to the ambient background noise. Frequency measurements for all three dolphin species namely bottlenose dolphin (*Tursiops truncates*), common dolphin (*Delphinus delphis*) and striped dolphin (*Stenella coeruleoalba*) have been generated using this modified spectrogram technique.

4.6. GM-PHD FILTER INITIALIZATION

The GM-PHD filter was initialized by considering the initial number of whistle frequencies as a random value. Then for each of these frequency components, a Gaussian component is constructed whose mean is distributed uniformly over the range 2 kHz-30 kHz [2]. This is because fundamental frequency of most whistles ranges from 2 to 30 kHz. The covariance matrix for each of each of the Gaussian components is chosen to be $\Sigma_v = diag\left(\left[\sigma_f^2 \; \sigma_\alpha^2\right]^T\right)$ where σ_f the standard deviation of frequency measurement noise and σ_α is the standard deviation of chirp rate noise. Experimentally, the best possible values for σ_f and σ_α are chosen to be 10 and 100 respectively. In essence initialization of the GM-PHD filter is the construction of a Gaussian mixture density made up with these individual Gaussian components. Target spawning is not relevant in whistle tracking and hence not considered.

4.7. SELECTION OF OTHER SIMULATION PARAMETERS

The clutter intensity κ_k is assumed to be independent of time, i.e. κ_k=κ. This is because the GM-PHD filter recursions are derived under such an assumption [12]. It is assumed that the measurement are generated using the modified spectrogram technique contains on the average $r = 5$ clutters which are uniformly distributed over the range of 2 kHz- 30 kHz and hence κ=$\frac{5}{28000}$. This produces best tracking performance in this application as shown in Fig.1-(a). The lowering of r adversary affect the GM-PHD filter output by including more clutter as illustrated in Fig.1-(b). If r is increased then filter starts to miss detection on tracks as in Fig.1-(c) and Fig.1-(d). A small value of w_{Th} causes the filter to consider Gaussian components with weak weights in the mixture density as outputs and this results in a more cluttered filter state estimate. Excessive large value for w_{Th} causes the filter to discard many Gaussian peaks that represents actual whistle track. This effect of w_{Th} on GM-PHD filter state estimation is shown in Fig.2 (a)-(d). It is evident from these results that w_{Th}=0.5 produces best tracking performance.
The GM-PHD recursion is derived assuming certain conditions are met [12]. One such assumption is that the target survival probability,p_S and the target detection probability, p_D are state independent. Hence the values of p_S and p_D are chosen to be 0.99 and 0.95 respectively in this application. With these values, the GM-PHD filter tracked all three tracks as illustrated in Fig.3-(a). Lowering values of p_S and $p_D$deteriorates the filter performance, i.e. the filter missed potential whistle tracks as in Fig.3-(b). The maximum number of Gaussian components J_{max} at each time step has been limited to 100 by using pruning method. The threshold for pruning, U and the truncation threshold, T_r have been chosen to be 4 and $1 \times 10^{-}5$ respectively. Table-1 summarizes the values used for different simulation parameters in the GM-PHD filter.

5. SIMULATION RESULTS FOR WHISTLE TRACKING USING GM-PHD FILTER

The GM-PHD filter produces the whistle frequency estimates based on the measurements generated from the hydrophone audio recording using modified spectrogram technique. The output of the GM-PHD filter for *Tursiops truncates* is superimposed on the spectrogram in Fig.4. For *Delphinus delphis* and *Stenella coeruleoalba*, the GM-PHD

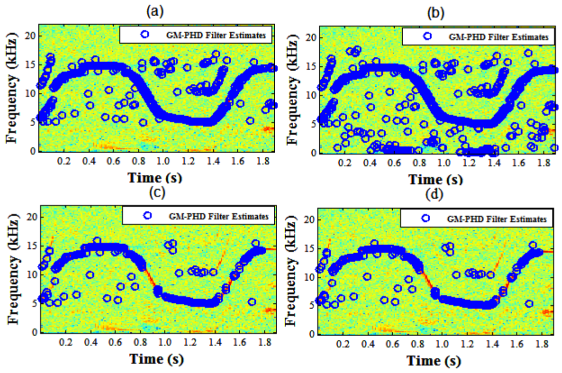

Figure 1: Estimated whistle tracks from the GM-PHD filter superimposed on spectrogram for *Tursiops truncates* for: (a) r=5, (b) r=2 (c) r=10, (d) r=15.

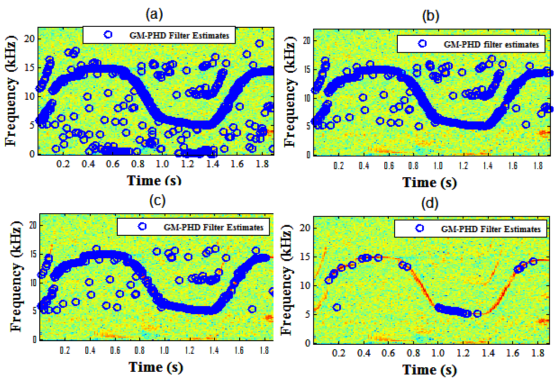

Figure 2: Estimated whistle tracks from the GM-PHD filter superimposed on spectrogram for *Tursiops truncates* for: (a) w_{Th}=0.2, (b) w_{Th}=0.5 (c) w_{Th}=0.7, (d) w_{Th}=0.9.

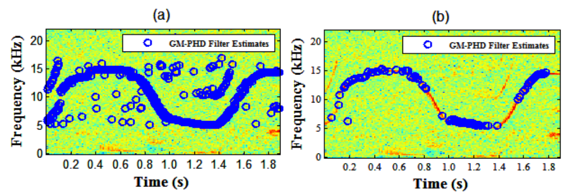

Figure 3: Estimated whistle tracks from the GM-PHD filter superimposed on spectrogram for *Tursiops truncates* for: (a) p_S=0.99 and p_D=0.95, (b) p_S=0.55 and p_D=0.50.

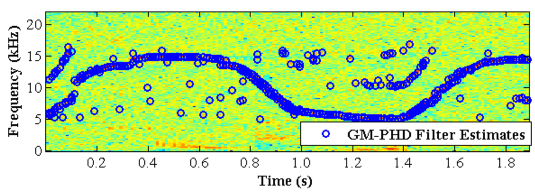

Figure 4: Estimated whistle tracks from the GM-PHD filter superimposed on spectrogram for *Tursiops truncates*.

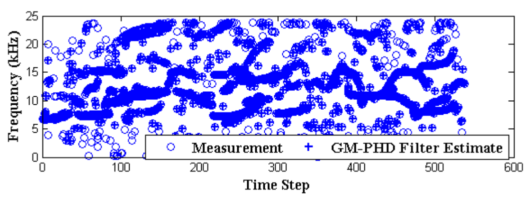

Figure 5: Estimated whistle tracks from the GM-PHD filter superimposed on measurements for *Delphinus delphis*.

Table 1: Summary of Simulation Parameters for the GM-PHD Filter in Dolphin Whistle Tracking.

Simulation Parameter	Parameter Specification
No. of Gaussian components for spontaneous birth, $J_{\gamma,k}$	$\|Z_k\|$
Average number of clutter point per scan, r	5
Standard Deviation of Frequency Measurement, σ_f	10
Standard Deviation of Sweep Rate Noise, σ_α	100
Covariance matrix for system noise, Σ_v	$diag\left(\left[\sigma_f^2\ \sigma_\alpha^2\right)\right]$
Initial Probability Distribution for frequency, f	$f \sim \mathcal{U}[2kHz, 30kHz]$
Initial Probability Distribution for sweep rate, α	$\alpha \sim \mathcal{U}[8kHz, 60kHz]$
Target Survival Probability, p_S	0.99
Target Detection Probability, p_D	0.95
Threshold for pruning, U	4
Maximum number of Gaussian Components allowed, J_{max}	100
Truncation threshold, T_r	1×10^{-5}
Intensity of Poisson RFS for Spontaneous birth, κ	$\frac{5}{28000}$
Weight threshold, w_{Th}	0.5
Sampling period, T	1 sec

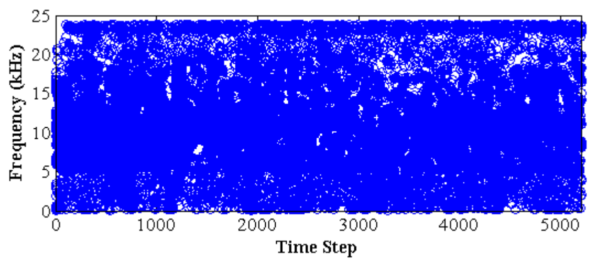

Figure 6: Estimated whistle tracks from the GM-PHD filter superimposed on measurements for *Stenella coeruleoalba*.

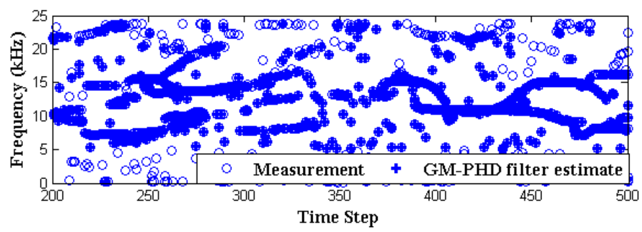

Figure 7: Estimated whistle tracks from the GM-PHD filter superimposed on measurements for *Delphinus delphis* expanded to show 300 time scan.

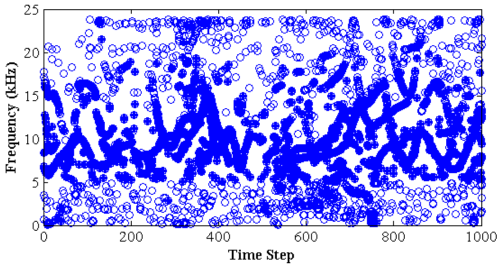

Figure 8: Estimated whistle tracks from the GM-PHD filter superimposed on measurements for *Stenella coeruleoalba* expanded to show 1000 time scan.

filter outputs are superimposed on the measurements in Fig.5 and Fig.6. For better pictorial representation, the GM-PHD filter outputs for *Delphinus delphis* (over 300 time scan) and *Stenella coeruleoalba* (over 1000 time scan) have been shown in Fig.7 and Fig.8 respectively. For all three species the simulation parameters, as specified in Table-1, are kept unchanged.

6. DISCUSSION ON RESULTS

Quantitative analysis of the whistle track estimation accuracy is not possible since there is no ground truth type track for the dolphin whistles. Ground truth tracks exist only in synthetic environment. Hence visual inspection of estimated tracks is used in dolphin whistle tracking community. Generally the spectrogram of the raw data, which consists of tracks and huge number of clutters, is visually inspected. This gives an idea about the possible whistle track positions and number. Now the goal of the GM-PHD filter is to produce track estimates on the possible tracks by discarding the clutters in the measurement. This has been achieved for bottlenose dolphin by adjusting the parameters of the GM-PHD filter as shown in Fig.1 to Fig.3. This automatic extraction of dolphin whistle tracks form the acoustic measurement is challenging due to different factors, as mentioned in Section 2, such as the presence of multiple whistles, multiple whistles crossing each other and number of whistles is varying over time. A close examination on the simulation results presented in Fig.4 to Fig.8 reveals the fact that the GM-PHD filter has successfully tracked multiple dolphin whistles from a set of noisy measurements for all three different dolphin species. The GM-PHD filter has tracked multiple whistles even when they cross each other as depicted in Fig.7 and Fig.8. At each time step the filter accurately calculated the expected number of whistles by calculating the number of local peaks in the Gaussian mixture target density and hence it can cope up with the time varying number of whistles. This allows the filter to produce correct number of estimates that will coincide with probable tracks and not with clutter. It also ensures that the tracks are automatically initiated and terminated as shown in Fig.7 and Fig.8. These results suggest that the automation in dolphin whistle tracking has been achieved using the GM-PHD filter.

7. Conclusions

This work has investigated the possibility of dolphin whistle track automation in the context MTT framework based on RFS. It has been demonstrated that satisfactory track estimates can be produced automatically using the GM-PHD filter. Simulation results suggest that the GM-PHD filter has successfully produced reliable track estimates in the presence of multiple whistles, spontaneous death/birth of whistles and multiple whistles crossing each other. In future these accurate track estimates produced by the GM-PHD filter can be combined with different dolphin species identification algorithms to build a complete system. The complete system will identify and classify different dolphin species by processing raw acoustic recordings obtained from hydrophones. Hence a complete dolphin species identification system based on PAM can be realized.

References

[1] J. N. Oswald, J. Barlow & T. F. Norris, Acoustic Identification of Nine Delphinid Species in the Eastern Tropical Pacific Ocean, 2003, *Marine Mammal Sci.*, 19(1), pp. 20-37.

[2] J. N. Oswald, S. Rankin, J. Barlow & M. O. Lammers , 2007, A Tool for Real-Time Acoustic Species Identification of Delphinid Whistles, *Journal of the Acoust. Soc. Am.*, 122(1), pp. 587-595.

[3] W. W. L. Au, *The Sonar of Dolphins*, 1993, Springer-Verlag, New York.

[4] , W. W. Steiner, 1981, Species-Specific Differences in Pure Tonal Whistle Vocalizations of Five Western North Atlantic Dolphin Species, *Behavioral Ecology and Sociobiology*, 9(4), pp. 241-246.

[5] T. A. Lampert & S. E. M. OâĂŹKeefe, A Survey of Spectrogram Track Detection Algorithms, 2010, *Applied Acoustics*, 71(2), pp. 87-100.

[6] R. Mahler, An Introduction to Multisource-Multitarget Statistics and Its Applications, 2000, *Lockheed Martin Technical Monograph*.

[7] V. J. Candy, 2009, Bayesian Signal Processing, John Wiley & Sons, New York.

[8] R. Mahler, 2007, *Statistical Multisource-Multitarget Information Fusion*, Artech House, New York.

[9] R. Mahler, Multi-Target Bayes Filtering via Frst-Order Multi-Target Moments, 2003, *IEEE Trans. on AES.*, 39(4), pp. 1152-1178.

[10] R. Mahler, "Statistics 101" for Multisensor, Multitarget Data Fusion, 2004, *IEEE Magazine on AES.*, 19(1), pp. 53-64.

[11] B. N. Vo, S. Singh & A. Doucet, Sequential Monte Carlo Methods for Multi-Target Filtering with Random Finite Sets, 2005, *IEEE Trans. on AES.*, 41(4), pp. 1224-1245.

[12] B. N. Vo & W. K. Ma, The Gaussian Mixture Probability Hypothesis Density Filter, 2006, *IEEE Trans. on SP.*, 54(11), pp. 4091-4104.

[13] J. Buck & P. Tyack, A Quantitative Measure Of Similarity for Tursiops Truncatus Signature Whistles, 1993, *Journal of the Acoust. Soc. Am.*, 94(5), pp. 2497-2506.

[14] S. Datta & C. Sturtivant, Dolphin Whistle Classification for Determining Group Identities, 2002, *Signal Process.*, 82(2), pp. 251-258.

[15] A. Mallawaarachchi, S. H. Ong, M. Chitre & E. Taylor, Spectrogram Denoising and Automated Extraction of the Fundamental Frequency Variation of Dolphin Whistles, 2008, *Journal of the Acoust. Soc. Am.*, 124(2), pp. 1159-1170.

[16] A. T. Johansson & P. R. White, An Adaptive Filter-Based Method for Robust, Automatic Detection and Frequency Estimation of Whistles, 2011, *Journal of the Acoust. Soc. Am.*, 130(2), pp. 893-903.

[17] K. Panta, Multi-Target Tracking using 1st Moment of Random Finite Sets, PhD Thesis, 2007 Department of EEE, The University of Melbourne, Australia.

[18] P. R. White & M. L. Hadley, Introduction To Particle Filters For Tracking Applications In The Passive Acoustic Monitoring Of Cetaceans, 2008, *Canadian Acous.*, 36(9), pp. 146-152.

[19] E. Maggio, M. Taj & A. Cavallaro, Efficient Multi-Target Visual Tracking Using Random Finite Sets, 2008, *IEEE Trans. on Circuits and Systems for Video Tech.*, 18(8), pp. 1016-1027.

[20] Y. Wang, J. Wu, A. A. Kassim & W. Huang, Data-Driven Probability Hypothesis Density Filter for Visual Tracking, 2008, *IEEE Trans. on Circuits and Systems for Video Tech.*, 18(8), pp. 1085-1095.

[21] D. Gillespie, P. R. White, M. Caillat & J. Gordon, Development and Implementation of Automatic Classification of Odontocetes within PAMGUARD, Workshop on Detection, Classification,Localization, and Density Estimation of Marine Mammals using Passive Acoustics Timberline Lodge, Mt. Hood, Oregon, USA.