Tracking Random Finite Objects using 3D-LIDAR in Marine Environments

Kwang Wee Lee, Bharath Kalyan,
Sardha Wijesoma, Martin Adams
Department of Electrical and Electronic Engg.
Nanyang Technological University
50 Nanyang Avenue, Singapore 639798
leekw,bharath, eswwijesoma,
EAdams{@ntu.edu.sg}

Franz S. Hover, Nicholas M. Patrikalakis
Department of Mechanical Engineering
Center for Ocean Engineering
Massachusetts Institute of Technology
Cambridge, MA 02139
hover, nmp{@mit.edu}

ABSTRACT

This paper presents a random finite set theoretic formulation for multi-object tracking as perceived by a 3D-LIDAR in a dynamic environment. It is mainly concerned with the joint detection and estimation of the unknown and time varying number of objects present in the environment and the dynamic state of these objects, given a set of measurements. This problem is particularly challenging in cluttered dynamic environments such as in urban settings or marine environments, because, given a measurement set, there is absolutely no knowledge of which object generated which measurement, and the detected measurements are indistinguishable from false alarms.

The proposed approach to multi-object tracking is based on the rigorous theory of finite set statistics (FISST). The optimal Bayesian multi-object tracking is not yet practical due to its computational complexity. However, a practical alternative to the optimal filter is the probability hypothesis density (PHD) filter, that propagates the first order statistical moment of the full multi-object posterior distribution. In contrast to classical approaches, this random finite set framework does not require any explicit data associations. In this paper, a Gaussian mixture approximation of the PHD filter is applied to track variable number of objects from 3D-LIDAR measurements by estimating both the number of objects and their respective locations in each scan. Experimental results obtained in marine environments demonstrate the efficacy and tracking performance of the proposed approach.

Keywords

Tracking, Velodyne, RFS, PHD filter

1. INTRODUCTION

One of the key tasks intelligent mobile vehicles have to perform is the reliable perception of their environment, namely the detection and tracking of multiple objects and free space. Laser scanners have proven to be efficient and less noisy in comparison with other ranging sensors such as radar and ultrasonic sensors that provide direct distance measurements. Most laser sensors reported in literature are restricted to two dimensions, that scan along a plane within a limited viewing angle. Each scan acquires a sequence of range and bearing measurements to objects within the planar sensing zone. This allows easy detection of objects in the environment by applying straight forward signal processing methods. However, objects above or below the scanning plane cannot be detected. The limited number of measurements thereby affects the accuracy of classification and tracking of objects. Additionally, in uneven terrains and marine environments where the sensor is subjected to pitch and roll due to vehicle motion, the scanner might fail to detect objects. Over the last few years, fully three-dimensional laser scanners have been introduced. These 3D laser scanners use an array of beams organized in multiple planes to provide range, bearing and azimuth data of objects. This allows detection of many kinds of objects and the explicit detection of free-space. However, the vast amount of data poses a great challenge on the processing algorithms.

The objective of multi-object tracking problem is to estimate the state of an unknown number of objects, based on the measurements of the objects corrupted by noise, in the presence of clutter. The classical approach for solving this problem is to use a stochastic filter such as Kalman filter or it's variants to each object and use a data association technique such as the nearest neighbor to assign the appropriate measurement to each filter and track each object separately[1], [10]. An alternative and a more elegant approach is to consider the multi-object set as a single meta object and the measurements received by the sensor as a single set of measurements [9] and modeling them as random finite sets (RFS). This allows estimating multiple objects in presence of clutter and with data association uncertainty to be cast in a Bayesian filtering framework.

The focus of this paper is on the application of the probability hypothesis density (PHD) filter, which is a recursion that propagates the first order statistical moment of the RFS of states in time, to track multiple objects in presence of measurement uncertainty and false alarms without any explicit data association. Due to its ability to handle non-linear and variable number of targets, it has been ap-

plied in various fields ranging from tracking multiple moving targets in uneven terrain [11] to detecting and tracking of underwater objects [6], [3]. Other notable applications of the PHD filter are in passive coherent location of targets observed from multiple bistatic radars [12], tracking corner features in optical image sequences [5] and tracking human figures in digital video [15].

Our main contribution in this paper is the adaptation of the PHD filter based on finite set statistics (FISST) to the complex real-world tracking scenarios using 3D scanning LIDAR, with particular emphasis on Velodyne HDL-64E. We demonstrate its performance with the data obtained from experiments conducted along the coastal waters of Singapore. The remainder of the paper is structured as follows. Section 2 details the data acquisition setup along with pre-processing steps involved. Section 3 reviews the basics of PHD filtering followed by the process and measurement models as used in the filter. It also discusses the Gaussian mixture (GM) PHD object tracking algorithm using 3D-LIDAR. Results, based on the experiments conducted using Velodyne HDL-64E along the coastal waters of Singapore are presented in Section 4. Section 5 concludes the paper.

2. DATA ACQUISITION & PROCESSING

Our method of object tracking is based on a scan-wise acquisition and processing which is performed in several steps (see Fig. 1): a scan acquisition from a Velodyne HDL-64E which is a 3D-LIDAR, followed by a segmentation and feature extraction, and finally a GM-PHD filter based object tracker. These steps are detailed in the following sections.
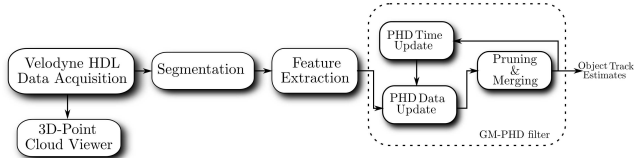

Figure 1: Block diagram of GM-PHD Object Tracker using 3D-LIDAR

2.1 Velodyne: A 3D-LIDAR

The Velodyne HDL-64E provides 3D range scans by rotating an array of 64 beams around its vertical axis at $5-15$ Hz (10 Hz in our application) and producing close to around 1.33 million points per second. In the horizontal direction, the array provides 360° field of view (FOV) with an angular resolution of approximately 0.09^o. Vertically, the pitch angles range from -24.8^o to $+2^o$ with an angular resolution of 0.4^o. Its range measurement accuracy typically is within 10 cm. The sensor is mounted on top of the mobile platform providing range scans with a full FOV in horizontal direction. A typical 360° range scan from the Velodyne HDL-64E is as shown in fig. 2.

2.2 Segmentation and Feature Extraction

The 3D point cloud data from each scan is projected onto a cylinder whose axis is the rotational axis of the LIDAR. This projection yields a range image, whose pixel intensity values correspond to the distance measurements as shown in fig. 3a. The bearing and azimuth index (u, v) in the range image is a direct mapping of the bearing and azimuth values (θ, ϕ) from the LIDAR, according to the following equation.

$$\left[\begin{array}{c} \theta \\ \phi \end{array} \right] = \left[\begin{array}{c} p_u u \\ p_{v_1} v + p_{v_2} v + p_{v_3} v + p_{v_4} v + p_{v_5} \end{array} \right] \quad (1)$$

where p_u, $p_{v_1} - p_{v_5}$ are the coefficients derived based on the polynomial curve fitting using the calibration parameters provided by the manufacturer. The range image can then be segmented using any standard range image segmentation method [4]. In this paper, we have used the mean shift segmentation technique [2] to segment objects in the range image. It mainly comprises of two steps: mean shift filtering of the original range image data, followed by clustering of the filtered data points. We then use the centroid of these segmented clusters as our measurement z to update the PHD filter, which is discussed in detail in the following section.

3. OBJECT TRACKING WITH PHD FILTER

This section describes the method for tracking multiple unknown number of objects from the 3D-LIDAR in the presence of false alarms (clutter). To achieve this, we use the PHD filter based on finite set statistics[8]. Modeling set valued states and measurements as RFS allows the problem of estimating multiple unknown of objects to be formulated in a multi-objective Bayesian filtering framework. However, the propagation of the full posterior distribution using the optimal multi-objective Bayesian approach is not practical due to computational complexity. A recursive Bayesian approach for approximating the first order statistical moment of the full posterior distribution known as the Probability Hypothesis Density (PHD) was proposed by [8] as a tractable alternative to the optimal multi-objective Bayes filter. However, the realization of the PHD filter involves multiple integrals that have no tractable closed form expressions in general. Sequential Monte-Carlo (SMC) [14] and Gaussian mixture (GM) [13] approximation techniques were devised to implement the PHD filter. In this paper, we apply the Gaussian mixture variant to implement the PHD filter for reliable tracking of the unknown and varying number of objects as observed by the 3D-LIDAR.

3.1 The PHD filter

Let the state of single object at time k be represented by $\mathrm{x}_k = \{\theta_k, \dot{\theta}_k, \phi_k, \dot{\phi}_k\} \in \mathcal{F}(\mathrm{x})$, where (θ_k, ϕ_k) are the object position and $(\dot{\theta}_k, \dot{\phi}_k)$ the object speed in the range image and $\mathcal{F}(\mathrm{x})$ is the single object space. Let the single object measurement at time k, which is as a result of segmentation and feature extraction from single LIDAR scan be represented by $\mathrm{z}_k = \{\theta_k, \phi_k\} \in \mathcal{F}(\mathrm{z})$ in the range image. Suppose at time k there are N_k objects and l_k measurements, then the corresponding multi-object states and the multi-object measurements are represented as finite sets $\mathrm{X}_k = \{\mathrm{x}_{k,1}, \ldots, \mathrm{x}_{k,N_k}\}$ and $\mathrm{Z}_k = \{\mathrm{z}_{k,1}, \ldots, \mathrm{z}_{k,l_k}\}$ which contain states of individual objects and measurements respectively [1]. The PHD filter recursion is a two step process:

- **PHD time update:** Given the process model, the

---

[1] $\mathrm{x}_{k,i}$ and $\mathrm{z}_{k,i}$ are denoted as x_k and z_k for notational simplicity.
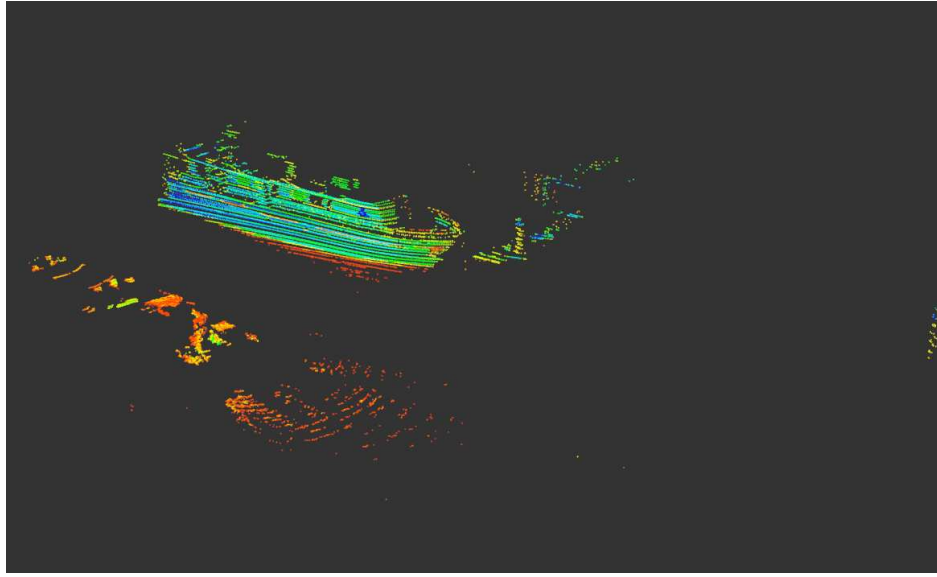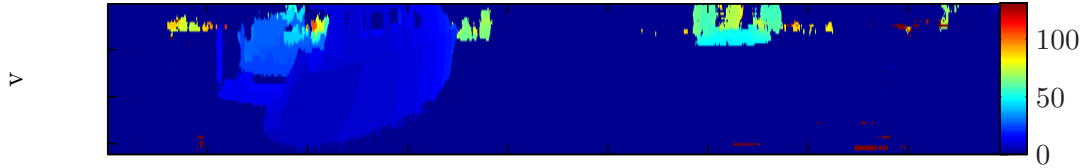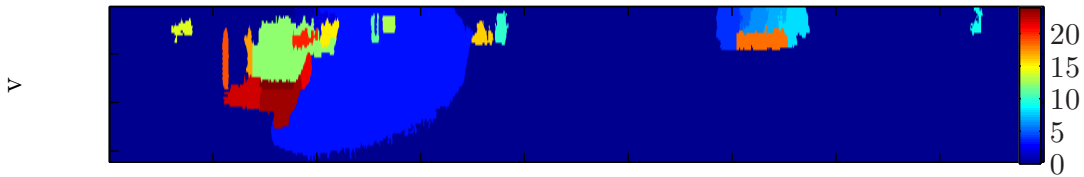
Figure 2: Velodyne HDL-64E scan represented as 3D point cloud. The intensity of the signals are color-mapped with darker colors representing stronger intensity returns.

(a) Range image. Each pixel value correspond to a distance measurement as indicated by the colorbar.

(b) Segmented range image. Each pixel value correspond to a segmented object as indicated by the colorbar.

Figure 3: Illustration of projection of point cloud from a scan in fig. 2 to obtain a range image. The result of mean-shift segmentation on the range image in (a) results in a segmented range image (b).

predicted PHD,

$$D_{k|k-1}(\mathbf{x}_k|\mathbf{Z}^{(k-1)}) = \overbrace{\gamma_k(\mathbf{x}_k)}^{\text{new objects}} + \int \overbrace{p_S(\mathbf{x}_{k-1}).f_{k|k-1}(\mathbf{x}_k|\mathbf{x}_{k-1})}^{\text{existing objects}}$$
$$.D_{k-1|k-1}(\mathbf{x}_{k-1}|\mathbf{Z}^{(k-1)})d\mathbf{x}_{k-1} \tag{2}$$

where,

- $\gamma_k(\mathbf{x}_k)$: PHD of the new incoming objects within the LIDAR field of view (FOV)
- $p_S(\mathbf{x}_{k-1})$: Probability of an object being re-observed

- **PHD data update:** Given a new set of measurements $\mathbf{Z}_k$, the updated PHD,

$$D_{k|k}(\mathbf{x}_k|\mathbf{Z}^{(k)}) = (1-p_D)D_{k|k-1}(\mathbf{x}_k|\mathbf{Z}^{(k-1)})$$
$$+ \sum_{\mathbf{z}_k \in \mathbf{Z}_k} \frac{p_D D_k(\mathbf{z}_k)}{\lambda_c c_k(\mathbf{z}_k) + p_D D_k(\mathbf{z}_k)} D_k(\mathbf{x}_k|\mathbf{z}_k) \tag{3}$$

where,

$$D_k(\mathbf{z}_k) = \int f_k(\mathbf{z}_k|\mathbf{x}_k)D_{k|k-1}(\mathbf{x}_k|\mathbf{Z}^{(k-1)})d\mathbf{x}_k \tag{4}$$

$$D_k(\mathbf{x}_k|\mathbf{z}_k) = \frac{f_k(\mathbf{z}_k|\mathbf{x}_k)D_{k|k-1}(\mathbf{x}_k|\mathbf{Z}^{(k-1)})}{D_k(\mathbf{z}_k)} \quad (5)$$

and,

- $f_k(\mathbf{z}_k|\mathbf{x}_k)$: is the sensor likelihood function $L_\mathbf{z}(\mathbf{x}_k)$
- λ_c: average number of false alarms per scan, which is assumed to be Poisson distributed
- $c_k(\mathbf{z}_k)$: distribution of each of the false alarms

3.2 Implementation of the GM-PHD filter tracker

In this work, we assume that each object moves according to the following linear Gaussian dynamics i.e.,

$$\mathbf{x}_k = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{k-1} + \begin{bmatrix} \frac{\Delta t^2}{2} & 0 \\ \Delta t & 0 \\ 0 & \frac{\Delta t^2}{2} \\ 0 & \Delta t \end{bmatrix} \begin{bmatrix} \mathbf{v}_{1,k-1} \\ \mathbf{v}_{2,k-1} \end{bmatrix}$$

$$(6)$$

and the measurement model

$$\mathbf{z}_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k \quad (7)$$

Thus, the state and the measurement process can be succinctly described as,

$$\mathbf{x}_k = F_k \mathbf{x}_{k-1} + G_k \mathbf{v}_{k-1} \quad (8)$$
$$\mathbf{z}_k = H_k \mathbf{x}_k + \mathbf{w}_k \quad (9)$$

where $\mathbf{v}_{k-1}$ and $\mathbf{w}_k$ are assumed to be zero mean Gaussian process noise and measurement noise with covariances Q_{k-1} and R_k, respectively. The implementation of the GM-PHD multi-object tracker is as proposed in [13]. For the benefit of the readers, we summarize the key steps of the 3D-LIDAR GM-PHD multi-object tracker in Table 1.

4. EXPERIMENTS AND RESULTS

In this section we report on the tests of the proposed multi-object tracking framework in real-world scenarios. In particular, we have evaluated the performance of the GM-PHD tracker for numerous scans acquired by the Velodyne HDL-64E sensor mounted on top of a research vessel (see fig. 4). As no ground truth information is available, a qualitative performance evaluation is conducted.

Figure 4: Research vessel used for the experiments.

Table 1 LIDAR GM-PHD Multi-object Tracker

● **Initialize**
At time $k = 0$, the PHD $D_{k|k}$ is initialized with a weighted sum of J_k Gaussians

$$D_{k|k}(\mathbf{x}|\mathbf{Z}_k) = \sum_{j=1}^{J_k} w_k^{[j]} \mathcal{N}(\mathbf{x}; \mu_k^{[j]}, \Sigma_k^{[j]})$$

These are distributed across the state space where each Gaussian term $\mathcal{N}(\mathbf{x}; \mu_k^{[j]}, \Sigma_k^{[j]})$ has a corresponding weight $w_k^{[j]}$, mean $\mu_k^{[j]}$, and variance $\Sigma_k^{[j]}$.
At $k \geq 1$,
●**Segmentation & Feature Detection**
The objects from the Velodyne HDL are detected using the segmentation and feature extraction techniques described in Section 2. The centroids of all the extracted blobs are the bearing-azimuth measurements represented by the set $\mathbf{Z}_k$ at time k.
● **PHD Time Update**
The predicted PHD up to time k is a Gaussian mixture,

$$D_{k|k-1}(\mathbf{x}) = D_{S,k|k-1}(\mathbf{x}) + \gamma_k(\mathbf{x})$$

where, $D_{S,k|k-1}(\mathbf{x})$ is predicted intensity of the existing (survived) objects in the FOV of the sensor, given by,

$$D_{S,k|k-1}(\mathbf{x}) = p_S \sum_{j=1}^{J_{k-1}} w_{k-1}^{[j]} \mathcal{N}(\mathbf{x}; \mu_{S,k|k-1}^{[j]}, \Sigma_{S,k|k-1}^{[j]})$$

with,

$$w_{k|k-1}^{[j]} = p_S . w_{k-1}^{[j]}; \quad \mu_{S,k|k-1}^{[j]} = F_{k-1}\mu_{k-1}^{[j]};$$
$$\Sigma_{S,k|k-1}^{[j]} = Q_{k-1} + F_{k-1}\Sigma_{k-1}^{[j]}F_{k-1}^T$$

and, $\gamma_k(\mathbf{x})$ is the PHD representing the new incoming objects in the FOV of the sensor, given by,

$$\gamma_k(\mathbf{x}) = \sum_{j=1}^{J_{\gamma,k}} w_{\gamma,k}^{[j]} \mathcal{N}(\mathbf{x}; \mu_{\gamma,k}^{[j]}, \Sigma_{\gamma,k}^{[j]})$$

with,

$$w_{k|k-1}^{[j]} = w_{\gamma,k}^{[j]}; \quad \mu_{k|k-1}^{[j]} = \mu_{\gamma,k}^{[j]}; \quad \Sigma_{k|k-1}^{[j]} = \Sigma_{\gamma,k}^{[j]}$$

● **PHD Data Update**
The PHD measurement update is a Gaussian mixture given by,

$$D_{k|k}(\mathbf{x}) = (1 - p_D)D_{k|k-1}(\mathbf{x}) + \sum_{z \in \mathbf{Z}_k} D_{L,k}(z|\mathbf{x})$$

where,

$$D_{L,k}(z|\mathbf{x}) = \sum_{j=1}^{J_{k-1}+J_{\gamma,k}} w_{k|k}^{[j]} \mathcal{N}(\mathbf{x}; \mu_{k|k}^{[j]}, \Sigma_{k|k}^{[j]})$$

with,

$$w_{k|k}^{[j]} = \frac{p_D w_{k|k-1}^{[j]} f_k^{[j]}(z|\mathbf{x})}{\lambda_c c_k(z) + \sum_{l=1}^{J_{k-1}+J_{\gamma,k}} w_{k|k-1}^{[l]} f_k^{[l]}(z|\mathbf{x})}$$

$$f_k^{[j]}(z|\mathbf{x}) = \mathcal{N}(z; H_k \mu_{k|k-1}^{[j]}, S_k^{[j]});$$

$$\mu_{k|k}^{[j]} = \mu_{k|k-1}^{[j]} + K_k^{[j]}[z - H_k\mu_{k|k-1}^{[j]}]; \Sigma_{k|k}^{[j]} = [I - K_k^{[j]}H_k^{[j]}]\Sigma_{k|k-1}^{[j]};$$

$$K_k^{[j]} = \Sigma_{k|k-1}^{[j]}[H_k^{[j]}]^T[S_k^{(j)}]^{-1}; S_k^{[j]} = R_k + H_k^{[j]}\Sigma_{k|k-1}^{[j]}[H_k^{[j]}]^T$$

Thus at time k, GM-PHD filter requires $J_k = (1 + |\mathbf{Z}_k|)(J_{k-1} + J_{\gamma,k})$ Gaussian components to represent the updated PHD with $(1 + |\mathbf{Z}_k|)$ components for each prediction term. The Gaussian mixture approximating the updated PHD is of the form,

$$D_{k|k}(\mathbf{x}) = \sum_{j=1}^{J_k} w_{k|k}^{[j]} \mathcal{N}(\mathbf{x}; \mu_k^{[j]}, \Sigma_k^{[j]})$$

● **Pruning & Merging**
In the pruning stage, the Gaussians with weights below a predetermined threshold τ_p, representing the updated PHD $D_{k|k}(\mathbf{x})$ are eliminated.
In the merging stage, the Gaussians whose distance between their means fall below a specific merging threshold τ_m, representing the updated PHD $D_{k|k}(\mathbf{x})$ are merged.
● **Object State Estimation**
The object states are obtained by selecting the Gaussians that are above a pre-determined threshold. In addition to these, the Gaussians that have already been classified as a valid object earlier are also included.
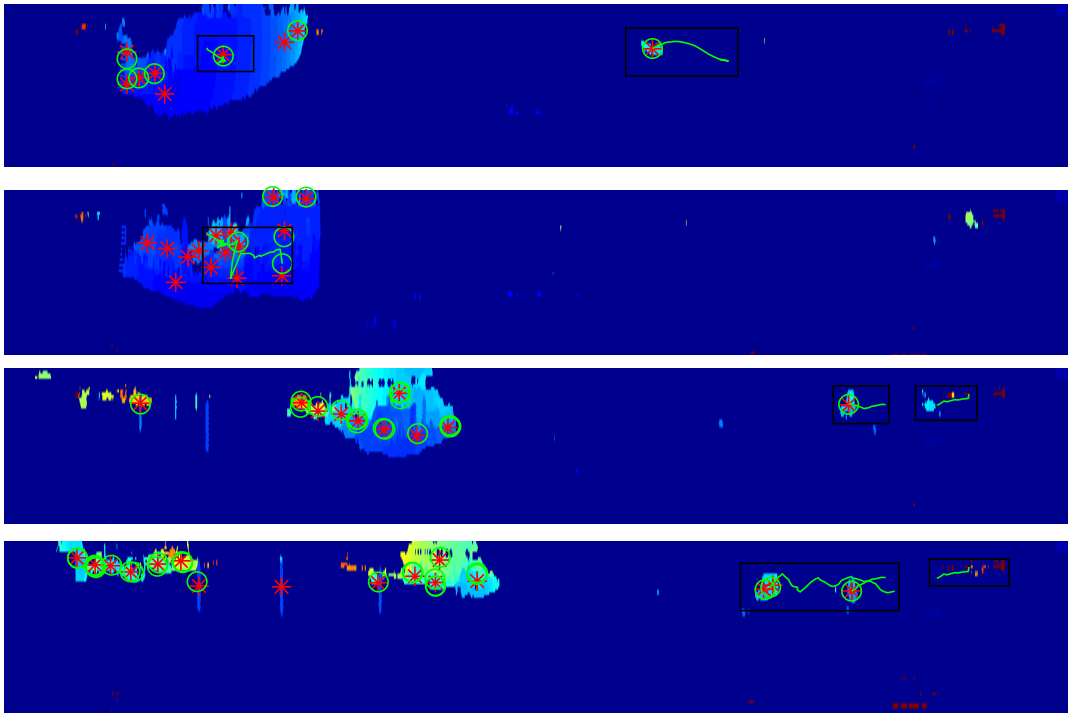
Figure 5: Illustration of tracking results from the GM-PHD filter. From top, frames 21, 49, 112 and 150. Asterix (red) are the cluster centroids, superimposed along with the filter output (green circles) and the tracks (green lines).

The parameters used for the multi-object tracker are as follows. The maximum number of Gaussian mixtures in the GM-PHD filter is limited to $J_k = 100$, with the Gaussian mixture pruning and merging parameters set to $\tau_p = 10^{-15}$ and $\tau_m = 4$ respectively. The pruning and merging process is necessary to reduce the number of Gaussian components propagated at each time step, which in turn reduces the computational complexity of the algorithm. These values chosen are based on a trade-off between the computational complexity and the quality of the tracking estimates.

From the position estimates of the objects (green circles) shown in fig. 5, it can be observed that the GM-PHD filter provides an accurate tracking performance. As noticed in frame number 150 in fig. 5, the filter accurately manages to detect and track the pitching motion of the ship. The PHD filter does generate false estimates at times (scan 21 in fig. 5). However, if the detections are not coherent and consecutive and the clutter is not persistent, then the PHD filter successfully manages to remove it.

5. CONCLUSION

We have presented a multi-object tracker that employs Gaussian mixture PHD filtering to remove clutter and missing detections from noisy measurements obtained from 3D LIDAR scans. The results demonstrate that the proposed algorithm successfully estimates and track the trajectories of the variable number of objects in dynamic marine environments.

The tracking case study presented here has a high SNR ratio, however it has been noted that under high cluttered environments and low SNR, PHD filter (as any other filter) performs poorly. To mitigate this problem, alternatives in form of cardinalized PHD (CPHD) filter [7] have been pro-

posed. Future work will assess the feasibility of applying CPHD filter to track in environments with higher clutter.

6. ACKNOWLEDGMENTS

7. REFERENCES

[1] Barshalom.Y and Fortman.T.E. *Tracking and Data Association*. Academic Press, 1988.

[2] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(5):603–619, 2002.

[3] D.E.Clark and J.Bell. Multi-target state estimation and track continuity for the particle PHD filter. *IEEE Transactions on Aerospace and Electronic Systems*, 43(3), July 2007.

[4] A. Hoover, G. Jean-Baptiste, X. Jiang, P. Flynn, H. Bunke, D. Goldgof, K. Bowyer, D. Eggert, A. Fitzgibbon, and R. Fisher. An experimental comparison of range image segmentation algorithms. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(7):689, 673, 1996.

[5] N. Ikoma, T. Uchino, and H. Maeda. Tracking of feature points in image sequence by SMC

implementation of PHD filter. In *Proc. Soc. of Instrument & Contr. Engineers (SICE)*, pages 1696–1071, Hokkaido, Japan, Aug. 2004.

[6] B. Kalyan, A. Balasuriya, and S. Wijesoma. Multiple target tracking in underwater sonar images using Particle-PHD filter. In *OCEANS 2006 - Asia Pacific*, pages 1–5, Singapore, 2006.

[7] R. Mahler. PHD filters of higher order in target number. *Aerospace and Electronic Systems, IEEE Transactions on*, 43(4):1523–1543, 2007.

[8] R. P. S. Mahler. Multi-target bayes filtering via first-order multi-target moments. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1152–1178, 2003.

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

[10] G. Pulford. Taxonomy of multiple target tracking methods. *Radar, Sonar and Navigation, IEE Proceedings -*, 152(5):291–304, 2005.

[11] H. Sidenbladh and S. Wirkander. Tracking random sets of vehicles in terrain. In *Proc. 2003 IEEE Workshop on Multi-Object Tracking*, Madison WI, June 2003.

[12] M. Tobias and A. D. Lanterman. Probability hypothesis density-based multitarget tracking with bistatic range and doppler observations. *IEE Proc. Radar, Sonar, and Navigation*, 152(3), 2005.

[13] B. Vo and W. Ma. The gaussian mixture probability hypothesis density filter. *IEEE Transactions on Signal Processing*, 54(11):4091–4104, Nov. 2006.

[14] B. Vo, S. Singh, and A. Doucet. Sequential monte carlo methods for multi-target filtering with random finite sets. *IEEE Transactions on Aerospace and Electronic Systems*, 41(4):1224–1245, 2005.

[15] Y. Wang, J. Wu, A. A. Kassim, and W. Huang. Tracking a variable number of human groups in video using probability hypothesis density. In *18th International Conference on Pattern Recognition*, Hong Kong, Aug. 2006.