Kayak Tracking using a Direct Lidar Model

Inger Berge Hagen
Department of Engineering Cybernetics
Norwegian University of Technology and Science
Trondheim, Norway
inger.b.hagen@ntnu.no

Edmund Brekke,
Department of Engineering Cybernetics
Norwegian University of Technology and Science
Trondheim, Norway
edmund.brekke@ntnu.no

***Abstract*—This paper proposes a direct approach for extended object tracking (EOT) using light detection and ranging (lidar) measurements. The method does not use any clustering operations, but processes the individual laser beams directly in an extended Kalman filter (EKF), and resolves data association by means of techniques reminiscent of the probabilistic data association filter (PDAF). The method is particularly tailored to tracking of kayaks, and parameterizes the shape of the kayak as a stick whose length is part of the state vector. The proposed method is evaluated through a simulation study and tested on real lidar data.**

I. Introduction

Autonomous surface vehicles (ASVs) have been getting increased attention within the maritime industry as they are potentially cost-saving and remove the need for placing human operators in hazardous environments and situations. However, the automation of surface vessels does come with its own set of challenges. On the one side there are technical challenges such as the demand for increased reliability and robustness of the vessel and its systems. Stable communication links are also a necessity, both for monitoring purposes and the possibility of remote control in case of emergencies. Other challenges are more directly related to the automation of the vessel itself, notably developing methods for collision avoidance and situational awareness.

To focus on the more automation related challenges it is desirable to concentrate on vessels traveling relatively short distances in coastal or urban waters where issues associated with communication and monitoring are greatly reduced and, if necessary, remote operation is an option. Urban ferries are therefore a natural choice as an experimental platform as they both follow a predefined itinerary within a restricted area and are generally close to existing infrastructures.

The challenges pertaining to autonomy will however not decrease. One example is the great variety in vessel types that may be encountered in such areas, anything from large barges to narrow and low-lying kayaks should be expected. Any ASV that operate in areas close to the shore should therefore have the ability to detect and track kayaks. The operating space can also be very restricted, e.g. a city canal. This combined with an often high traffic density will leave little space for evasive maneuvers and safety margins. Having accurate knowledge of both the position and extent of surrounding vessels is therefore crucial to safe operations.

It is also imperative that any collision avoidance maneuvers are made in a predictable and safe manner. It follows that information pertaining to the speed and heading of other vessels, along with corresponding uncertainty measures, is necessary for the planning of collision free trajectories that uphold acceptable safety margins.

The information needed by the tracking system is obtained via sensors, the most common in use for marine applications being the marine radar. In conventional tracking systems it is assumed that a target generates at most one measurement per scan. In cases where multiple returns are received, clustering techniques are used to join measurements that are likely to originate from the same target. Relevant information about target extent can thus be lost.

To retain information about target extent a high-resolution sensor, yielding multiple measurements per target, can be combined with an EOT [1] approach where the extent of a target is included in its estimated state vector. The lidar sensor is an example of such a high-resolution sensor which in combination with different EOT approaches has been employed for the tracking of cars [2], pedestrians [3], [4] and boats [5], [6].

The ellipse is a common model for marine vessels and is used in both [5] and [6]. The former [5] presents the generalized probabilistic data assosiation (GPDA) filter which combines a random matrix approach with probabilistic data association. In the latter [6], the random matrix approach was replaced with the contour tracking from [7] and the two methods were compared, both in simulations and on real lidar data. In both cases contour tracking outperformed the random matrix approach, motivating further research into more direct methods where the inherent structure of the measurements is exploited.

While an elliptic approximation is suitable for many marine vessels, the highly structured lidar measurements will for the particular case of a kayak target, produce returns along a very narrow shape, see Fig. 1. It then becomes questionable whether one can estimate the parameters (short and long axes) of a full ellipse, and a simpler stick model with only the long axis retained is a viable alternative. This simple stick approximation also maintains the inherent structure in the lidar measurements.

We have investigated the feasibility of tracking a kayak by means of lidar, and we present a singe-target EOT method

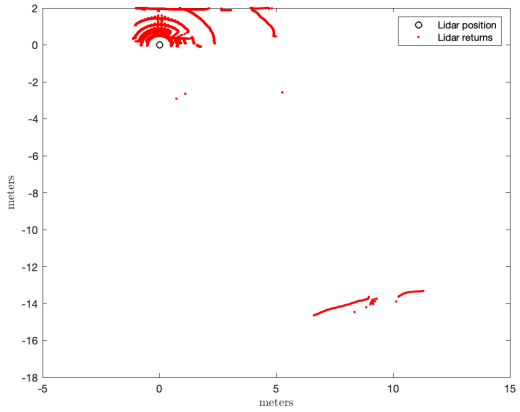


Fig. 1: Lidar measurements of a kayak projected on to a plane. The returns from the kayak forms the line in the bottom right corner.

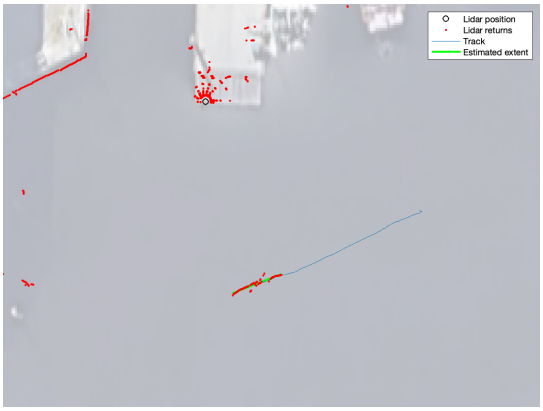


Fig. 2: Lidar data with track on kayak and its estimated extent.

tailored to the stick approximation. The data association is based on principles and assumptions similar to the ones underlying the PDAF [8]. A key challenge in EOT is to sample an appropriately diverse set of data association hypotheses. We explore sampling techniques based on random sample consensus (RANSAC), a brief description of which can be found in [9]. RANSAC was also the basis for the tracking methods presented in [10] and [11].

II. Problem Formulation

Our aim being to demonstrate the feasibility of using a direct measurement model, the problem has been limited to tracking a target located within the range of a stationary lidar sensor and issues such as target birth and death has been left for later work.

A. State model

The state vector at time step k is denoted by $\boldsymbol{x}_k = [x_k, y_k, u_k, v_k, \phi_k, l_k]$, where (x_k, y_k) signifies the position of the target's centroid in Cartesian coordinates, (u_k, v_k) the velocities, ϕ_k the heading and l_k the length of the target. The dynamic evolution of the target state is modeled by a linear Gaussian model

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{k-1} + \boldsymbol{q}_k, \quad p(\boldsymbol{q}_k) = \mathcal{N}(\boldsymbol{q}_k; \boldsymbol{0}, \boldsymbol{Q}), \tag{1}$$

where the matrices are given by the discrete-time constant velocity model

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{I}_2 & \boldsymbol{A} \\ \boldsymbol{0}_4 & \boldsymbol{I}_4 \end{bmatrix}, \quad \boldsymbol{A} = \begin{bmatrix} T_s\boldsymbol{I}_2 & \boldsymbol{0}_2 \end{bmatrix}, \tag{2}$$

$$\boldsymbol{Q} = \begin{bmatrix} \boldsymbol{B}\otimes\boldsymbol{I}_2 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C} \end{bmatrix}, \ \boldsymbol{B} = \sigma_p^2 \begin{bmatrix} \frac{T_s^3}{3} & \frac{T_s^2}{2} \\ \frac{T_s^2}{2} & T_s \end{bmatrix}, \ \boldsymbol{C} = \begin{bmatrix} \sigma_a^2 T_s & 0 \\ 0 & \sigma_l^2 \end{bmatrix}, \tag{3}$$

where $\boldsymbol{I}_n$ is the $n \times n$ identity matrix and $\boldsymbol{0}_m$ is a $m \times m$ zero matrix. The sample time is denoted T_s and the noise covariance for the acceleration and length is given by σ_a^2 and σ_l respectively.

B. Measurement model

Lidar sensors calculate distances by emitting laser light and measuring the reflected light. The 3D position of a return relative to the sensor is thus given by the measured distance along with the horizontal and vertical angle of the laser beam which together form a measurement, denoted $\boldsymbol{z} = [r \ \theta]^\intercal$. In the following sections it is assumed that the sensor is a side-looking 2D lidar emitting laser beams at fixed angular intervals.

The radial component r_k^j for a measurement will be within the interval $[0, R_{max}]$, where R_{max} is the maximum range of the sensor. The angular component is given by the angle of the beam that contains the detection. The set of candidate measurements for target association is selected by forming an elliptical validation gate around the predicted target extent. The area and orientation of the gate is decided by mapping the unit circle onto an ellipse by the transformation matrix:

$$\boldsymbol{T} = \boldsymbol{S}\boldsymbol{R}^\intercal + \sqrt{\boldsymbol{L}}\boldsymbol{V}^\intercal, \tag{4}$$

where the first term accounts for the predicted extent $\bar{l}_k$ and heading $\bar{\phi}_k$. It consist of the rotation matrix $\boldsymbol{R}$ and the scaling matrix $\boldsymbol{S}$.

$$\boldsymbol{R} = \begin{bmatrix} \cos\bar{\phi}_k & -\sin\bar{\phi}_k \\ \sin\bar{\phi}_k & \cos\bar{\phi}_k \end{bmatrix} \ \boldsymbol{S} = \begin{bmatrix} \bar{l}_k & 0 \\ 0 & 1 \end{bmatrix}, \tag{5}$$

The second term is the contribution from the covariance of the position, where $\boldsymbol{V}$ is a matrix whose columns are the eigenvectors of $\boldsymbol{Q}_{1:2,1:2}$ and $\boldsymbol{L}$ is the diagonal matrix whose non-zero elements are the corresponding eigenvalues. As measurements can only appear on the beams the volume of the validation region V_k is equal to the length of the beams within this region.

The list of target generated measurements falling within the gate is denoted $Z_{T,k} = \{\boldsymbol{z}_k^j\}_{j=1}^{\nu_k}$, the list of clutter measurements $Z_{C,k} = \{\boldsymbol{z}_k^j\}_{j=1}^{\mu_k}$, ν_k and μ_k denote the number of measurements. The list

$$Z_k = \{\boldsymbol{z}_k^j\}_{j=1}^{m_k}, \quad m_k = \mu_k + \nu_k \tag{6}$$

contains all validated measurements collected at time step k. Sensor properties assure that Z_k is ordered according to the angular component of the measurements so that $\theta_k^j < \theta_k^{j+1}$.

The measurement likelihood for the list of measurements can then be expressed as:

$$p(Z_k|\boldsymbol{x}_k) = \prod_{\boldsymbol{z}_k^j \in Z_{C,k}} p^c(\boldsymbol{z}_k^j) \prod_{\boldsymbol{z}_k^j \in Z_{T,k}} p^d(\boldsymbol{z}_k^j|\boldsymbol{x}_k). \quad (7)$$

Each factor in the above expression can be seen as a mixture:

$$p^c(\boldsymbol{z}_k^j) = P_{FA} p^{c1}(\boldsymbol{z}_k^j) + (1 - P_{FA}) p^{c0}(\boldsymbol{z}_k^j) \quad (8a)$$
$$p^d(\boldsymbol{z}_j^j) = P_D p^{d1}(\boldsymbol{z}_k^j) + (1 - P_D) p^{d0}(\boldsymbol{z}_k^j)., \quad (8b)$$

where P_{FA} denotes the probability of false alarms and P_D the number of detections. The first term in Equation 8a denotes the clutter distribution, the first term in 8b the measurement likelihood for a single measurement. It should be noted that while the expressions in Equation 8 can be seen as a virtual model, they do however give a mathematical correct representation.

C. Hypothesis Generation

When a list of measurements Z_k is received, a random sample consensus (RANSAC) method is employed to generate n_k linear models, fitted to the measurements. The method works by drawing a minumum sample set (MSS) from the measurements, then fitting a linear model to the MSS. The compatibility of the remaining measurements with the model is tested by calculating the euclidean distance to the line and measurements within a certain distance are added to the hypothesis. This process is repeated for a fixed number of iterations N. The resulting hypotheses are ranked using a scoring function and those with scores below a chosen threshold is rejected such that $n_k \leq N$ hypotheses are retained, each separating the measurements into two sets

$$\begin{aligned} \mathcal{D}_k^i &= \{\boldsymbol{z} \in Z_k; \boldsymbol{z} \text{ is an inlier of model } i\} \\ \mathcal{C}_k^i &= \{\boldsymbol{z} \in Z_k; \boldsymbol{z} \notin \mathcal{D}_k^i\}, \end{aligned} \quad (9)$$

where the measurements contained in $\mathcal{D}^i$ are considered target originated (inliers), and those in $\mathcal{C}^i$ are considered clutter (outliers).

The hypothesis should also include misdetections, to achieve this we define the set $\mathcal{G}_k$ as a set containing one measurement for each beam that falls within the gate at time k. For beams that does not contain a real measurement, the radial component is set to the hyper maximal range, denoted R_{hm} as to not confuse them with any real measurements. This allows the introduction of the following sets:

$$\begin{aligned} \mathcal{T}_k^i &= \{\boldsymbol{z} : \boldsymbol{z} \in \mathcal{G}_k \wedge min_\theta(\mathcal{D}^i) \leq \theta \leq max_\theta(\mathcal{D}^i)\} \\ \mathcal{M}_k^i &= \mathcal{T}_k^i \setminus \mathcal{D} \end{aligned} \quad (10)$$

where $\mathcal{T}_k^i$ is the set of beams that intersect with the target and $\mathcal{M}_k^i$ is the set of beams with misdetections. [1]

Together, the sets $\mathcal{D}^i$, $\mathcal{C}^i$ and $\mathcal{M}^i$ form an association hypothesis $a_k^i = \{\mathcal{D}^i, \mathcal{C}^i, \mathcal{M}^i\}$. The set of all association hypotheses generated at time k is $\mathcal{A}_k = \{a_k^i\}_{i=1}^{n_k}$.

[1]The set $\mathcal{T}_k^i$ can be extended to form additional hypotheses assuming misdetections at the target's endpoints.

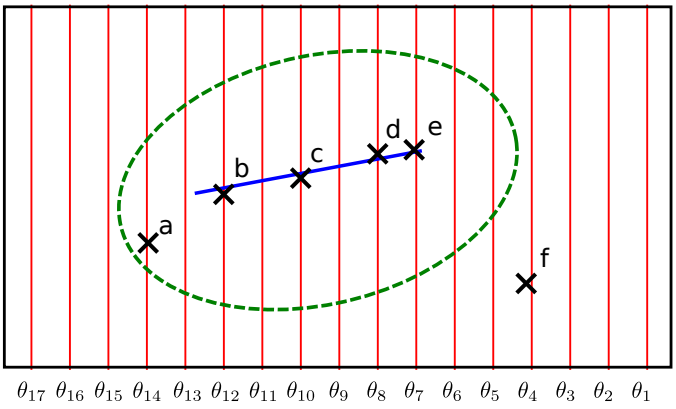


Fig. 3: Sets generated generated under the hypothesis that measurements b, c, d and e are target originated, gate marked with (- -). Inliers $\mathcal{D} = \{\boldsymbol{z}_b, \boldsymbol{z}_c, \boldsymbol{z}_d, \boldsymbol{z}_e\}$, outliers $\mathcal{C} = \{\boldsymbol{z}_a\}$, measurements from all gated beams $\mathcal{G} = \{\boldsymbol{z}_{\theta_j}\}_{j=5}^{14}$ where $\boldsymbol{z}_{\theta_j} = [R_{hm} \ \theta_j]^\intercal$, target intersecting beam measurements $\mathcal{T} = \{\boldsymbol{z}_{\theta_j}\}_{j=57}^{12}$ and misdetections $\mathcal{M} = \{\boldsymbol{z}_{\theta_9} \ \boldsymbol{z}_{\theta_{11}}\}$.

D. Hypotheses dependent innovations

A key difference between standard single-tracking and EOT is that in the context of EOT the innovation, i.e. the difference between prediction and measurements, must be calculated for multiple measurements per target. In general there will be a difference between the set of beams touched by the predicted kayak and the beams touched by the real kayak. To be able deal with this discrepancy within the EKF/PDAF framework the concept of virtual measurements is introduced. This approach is explained in the following paragraphs, illustrated by an example in Figure 4.

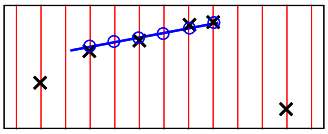
1) Prediction: Given a predicted target state $\bar{\boldsymbol{x}}_k$, intersection points between the lidar beams and predicted target extent is calculated. Each point is then converted into polar coordinates according to the following nonlinear equation

$$\begin{aligned} \boldsymbol{z} &= [r \ \theta]^\intercal \\ &= \left[\sqrt{x^2 + y^2} \ \ \text{atan2}(y, x)\right]^\intercal. \end{aligned} \quad (11)$$

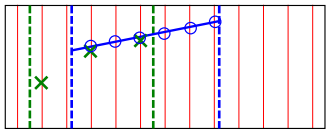
This yields a list of predicted measurements $\bar{Z}_k'$ where each element is on the form described in Eq. 11. A visualized example is shown in Figure 4a.

2) Endpoint Angle Calculation: The definition of the association hypotheses, Section II-C can be utilized to approximate the extent of the target. For a given association hypothesis a_k^i, a rough estimate of the endpoints is half-way between a beam covered by the target and a beam not covered by the target. The lower and upper angles of the hypothesized target extent are denoted $\hat{\theta}_k^l$ and $\hat{\theta}_k^u$ respectively. For the predicted extent, the endpoint angles $\bar{\theta}_k^l$ and $\bar{\theta}_k^u$ are calculated from the state vector. All angles are illustrated in Figure 4b.

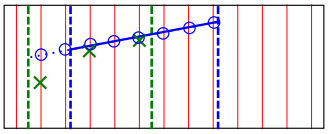
3) Extension: To compare predictions and actual measurements that lie on the same beam, the length parameter l_k of the predicted state $\bar{\boldsymbol{x}}_k$ is extended to cover beams that have target originated measurements, shown in Figure 4c.



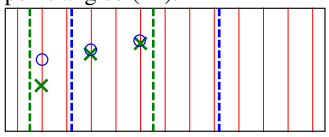
(a) Received measurements (×), predicted extent (—) and predicted measurements (○).



(b) Hypothesized target measurements (×) and endpoint angles (- -) along with predicted endpoint angles (- -).



(c) Extending the predicted target extent (⋯) resulting in two additional measurement predictions.



(d) Removal of predictions that does not align with the current hypothesis.

Fig. 4: Illustration of predicted measurements as a function of the predicted state and a given association event.

4) Removal: The final step is to remove any predictions that do not align with real measurements. The list of predictions $\bar{Z}'_k$ and the set target originated measurements $\mathcal{D}^i_k$ are now of equal size, as shown in Figure 4c.

5) Restructuring: The angular components of the remaining measurements are deterministically given and can therefore be removed. To amend for the information loss in the extension and removal step, the endpoint angles are then included. This restructuring is performed on both $\bar{Z}'_k$ and $\mathcal{D}^i_k$ resulting in two virtual measurement lists $\bar{Z}^i_k$ (predictions) and $\hat{Z}^i_k$ (target measurement). Both lists are on the form $Z = [\boldsymbol{r}^\intercal,\ \theta^l,\ \theta^u]^\intercal$, where $\boldsymbol{r}$ is a column vector containing the radial component of the remaining measurements.

The above previously described steps are performed in the nonlinear measurement function

$$\bar{Z}^i_k = \boldsymbol{h}_i(\bar{\boldsymbol{x}}_k). \tag{12}$$

The Jacobian of the hypothesis conditioned measurement function is found by finite difference, linearizing around the predicted state vector $\bar{\boldsymbol{x}}_k$ and hypothesized target measurements $\hat{Z}^{i\prime}_k$.

$$\boldsymbol{H}^i_k = \frac{\mathrm{d}\boldsymbol{h}}{\mathrm{d}\boldsymbol{x}}\Big|_{\bar{\boldsymbol{x}}_k, \hat{Z}^{i\prime}_k} \tag{13}$$

The covariance for each of the radial measurements is σ_r^2, these are collected in the vector $\boldsymbol{\sigma}_r$. The distributions of the two endpoint angles are approximated by Gaussians giving the covariance matrix $\boldsymbol{R}^i_k$.

$$\boldsymbol{A} = \mathrm{diag}(\boldsymbol{\sigma}_r) \quad \boldsymbol{B} = \begin{bmatrix} \frac{\sigma_\theta^2}{12} & 0 \\ 0 & \frac{\sigma_\theta^2}{12} \end{bmatrix} \quad \boldsymbol{R}^i_k = \begin{bmatrix} \boldsymbol{A} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{B} \end{bmatrix} \tag{14}$$

III. Tracking approach

According to the total probability theorem the posterior density of the target can be written as

$$p_k(\boldsymbol{x}_k) = \sum_{\mathcal{A}_k} p(\boldsymbol{x}_k | a^i_k, Z_{1:k}) \mathrm{P}(a^i_k | Z_{1:k}) \tag{15}$$

where $Z_{1:k} = \{Z_l\}_{l=1}^k$ is the cumulative set of measurements at time step k.

A. State and Covariance Update

The event-conditional densities are given by

$$\begin{aligned} p(\boldsymbol{x}_k | a^i_k, Z_{1:k}) &\propto p(Z_k | \boldsymbol{x}_k, a^i_k) p(\boldsymbol{x}_k | Z_{1:k-1}) \\ &= \prod_{\boldsymbol{z}^j_k \in \mathcal{C}^i} p^c(\boldsymbol{z}^j_k) \prod_{\boldsymbol{z}^j_k \in \mathcal{D}^i} p(\boldsymbol{z}^j_k | \boldsymbol{x}_k)\ p(\boldsymbol{x}_k | Z_{1:k-1}) \\ &= \frac{1}{V_k^{\phi^i_k}} \mathcal{N}(\hat{Z}^i_k; \boldsymbol{h}_i(\bar{\boldsymbol{x}}_k), \boldsymbol{R}^i_k) \mathcal{N}(\boldsymbol{x}_k; \bar{\boldsymbol{x}}_k, \bar{\boldsymbol{P}}_k) \\ &\propto \mathcal{N}(\boldsymbol{x}_k; \hat{\boldsymbol{x}}^i_k, \hat{\boldsymbol{P}}^i_k) \end{aligned} \tag{16}$$

where V_k is the total beam length within the gate. This results in the following expressions for the filter:

$$\begin{aligned} \boldsymbol{S}^i_k &= \boldsymbol{H}^i_k \bar{\boldsymbol{P}}_k \boldsymbol{H}^{i\intercal}_k + \boldsymbol{R}^i_k & \boldsymbol{W}^i_k &= \bar{\boldsymbol{P}}_k \boldsymbol{H}^{i\intercal}_k \boldsymbol{S}^{i\,-1}_k \\ \boldsymbol{\nu}^i_k &= \hat{\boldsymbol{Z}}^i_k - \boldsymbol{h}_i(\bar{\boldsymbol{x}}_k) & \hat{\boldsymbol{x}}^i_k &= \bar{\boldsymbol{x}}_k + \boldsymbol{W}^i_k \boldsymbol{\nu}^i_k \\ \hat{\boldsymbol{P}}^i_k &= (\boldsymbol{H}^{i\intercal}_k \boldsymbol{R}^{i\,-1}_k \boldsymbol{H}^i_k + \bar{\boldsymbol{P}}_k^{-1})^{-1} \end{aligned}$$

The covariance of the updated state is

$$\hat{\boldsymbol{P}}_k = \boldsymbol{P}^c_k + \tilde{\boldsymbol{P}}_k \tag{17}$$

where the spread of innovation term is given by

$$\begin{aligned} \tilde{\boldsymbol{P}}_k = \sum_{i=1}^{n_k} \beta^i_k \left[\bar{\boldsymbol{x}}_k + \boldsymbol{K}^i_k \boldsymbol{\nu}^i_k\right] \left[\bar{\boldsymbol{x}} + \boldsymbol{\nu}^i_k \boldsymbol{K}^i_k\right]^\intercal \\ - \left[\bar{\boldsymbol{x}}_k + \boldsymbol{K}^i_k \boldsymbol{\nu}^i_k\right] \left[\bar{\boldsymbol{x}} + \boldsymbol{\nu}^i_k \boldsymbol{K}^i_k\right]^\intercal \end{aligned} \tag{18}$$

where $\beta^i_k \triangleq \mathrm{P}\{a^i_k\}$. The covariance of the state updated with the correct measurement is

$$\boldsymbol{P}^c_k = \sum_{i=1}^{n_k} \beta^i_k (\boldsymbol{H}^{i\intercal}_k \boldsymbol{R}^{i\,-1}_k \boldsymbol{H}^i_k + \bar{\boldsymbol{P}}_k^{-1})^{-1}. \tag{19}$$

The Equations (19) and (18) are derived in a similar manner to the expressions used in the PDAF [8], but without a zero hypothesis.

B. Association Probabilities

The measurement likelihood for a given hypothesis is given by

$$p(Z_k | \boldsymbol{x}_k) = \prod_{\boldsymbol{z}^j_k \in \mathcal{N}^i_k} p^c(\boldsymbol{z}^j_k) \prod_{\boldsymbol{z}^j_k \in \mathcal{T}^i} p^d(\boldsymbol{z}^j_k | \boldsymbol{x}_k), \tag{20}$$

where $\mathcal{N}^i_k = \mathcal{G}_k \setminus \mathcal{T}^i_k$, i.e. measurements that fall within the gate, but are not associated with the target under the i^{th} hypothesis. As described in Section II-B each factor in the above expression can be seen as a mixture:

$$p^c(\boldsymbol{z}^j_k) = P_{FA} p^{c1}(\boldsymbol{z}^j_k) + (1 - P_{FA}) p^{c0}(\boldsymbol{z}^j_k) \tag{8a}$$

$$p^d(\boldsymbol{z}^j_j) = P_D p^{d1}(\boldsymbol{z}^j_k) + (1 - P_D) p^{d0}(\boldsymbol{z}^j_k). \tag{8b}$$

The first term in each of the expressions in Equation 8 denotes the clutter distribution and measurement likelihood respectively. The last term in each represent the likelihood

of a missed clutter measurement and a missed detection respectively and can be defined as

$$p^{c0}(\boldsymbol{z}_k^j) \triangleq \delta(r_k^j - R_{hm}) \tag{22a}$$

$$p^{d0}(\boldsymbol{z}_k^j) \triangleq \delta(r_k^j - R_{hm}), \tag{22b}$$

where $\delta(\cdot)$ is the Dirac delta function and r_k^j is the radial component of measurement $\boldsymbol{z}_k^j$. This can be viewed as a virtual model as the important property being that they are equal and can be canceled against each other. With this in mind and using $c_1 = \left(\frac{1-P_D}{1-P_{FA}}\right)$ and $c_2 = \left(\frac{P_D V_k}{P_{FA}}\right)$, the association probabilities can be calculated as follows:

$$\begin{aligned}
\mathrm{P}\{a_k^i\} &\propto \frac{1}{\prod_{\boldsymbol{z}_k^j \in \mathcal{G}_k} p^c(\boldsymbol{z}_k^j)} \prod_{\boldsymbol{z}_k^j \in \mathcal{N}^i} p^c(\boldsymbol{z}_k^j) \int \left(\prod_{\boldsymbol{z}_k^j \in \mathcal{T}^i} p^d(\boldsymbol{z}_k^j | \boldsymbol{x}_k) \right) p(\boldsymbol{x}_k) \mathrm{d}\boldsymbol{x}_k \\
&= \int \prod_{\boldsymbol{z}_k^j \in \mathcal{T}^i} \frac{p^d(\boldsymbol{z}_k^j | \boldsymbol{x}_k)}{p^c(\boldsymbol{z}_k^j)} p(\boldsymbol{x}_k)) \mathrm{d}\boldsymbol{x}_k \\
&= \int \left(\prod_{\boldsymbol{z}_k^j \in \mathcal{D}^i} \frac{P_D V_k}{P_{FA}} \mathcal{N}(\boldsymbol{z}_k^j; \boldsymbol{h}^{(j)}(\boldsymbol{x}_k), \sigma_r^2) \right) c_1^{\rho_k^i} p(\boldsymbol{x}_k) \mathrm{d}\boldsymbol{x}_k \\
&= c_1^{\rho_k^i} c_2^{\nu_k^i} \int \mathcal{N}(\hat{Z}_k^i; \boldsymbol{h}_i(\bar{\boldsymbol{x}}_k), \boldsymbol{R}_k^i) p(\boldsymbol{x}_k) \mathrm{d}\boldsymbol{x}_k \\
&= c_1^{\rho_k^i} c_2^{\nu_k^i} \int \mathcal{N}(\hat{Z}_k^i; \boldsymbol{h}_i(\bar{\boldsymbol{x}}_k), \boldsymbol{H}_k^i \bar{\boldsymbol{P}}_k \boldsymbol{H}_k^{i\intercal} + \boldsymbol{R}_k^i) \mathcal{N}(\boldsymbol{x}_k; \bar{\boldsymbol{x}}_k, \bar{\boldsymbol{P}}_k) \mathrm{d}\boldsymbol{x}_k \\
&= c_1^{\rho_k^i} c_2^{\nu_k^i} \mathcal{N}(\hat{Z}_k^i; \boldsymbol{h}_i(\bar{\boldsymbol{x}}_k), \boldsymbol{H}_k^i \bar{\boldsymbol{P}}_k \boldsymbol{H}_k^{i\intercal} + \boldsymbol{R}_k^i),
\end{aligned} \tag{23}$$

where ρ_k^i is the number of misdetections and ν_k^i is the number of detections.

IV. Simulation Study

A simulation study of was performed using as input the simulated output of a laser scanner with a range of 100 meters and 360° field of view. The initial surge velocity of the target was uniformly drawn from the interval [0.3m/s, 0.6m/s]. Initial values for position and heading variables was uniformly drawn from the intervals shown in Table I. The intervals were chosen so that the target was likely to stay within range throughout the simulation.

Quadrant	x interval	y interval	ϕ interval
1	[40m, 60m]	[40m, 60m]	[170°, 190°]
			[260°, 280°]
2	[-40m, -60m]	[40m, 60m]	[−10°, 10°]
			[260°, 280°]
3	[-40m, -60m]	[-40m, -60m]	[−10°, 10°]
			[80°, 100°]
4	[40m, 60m]	[-40m, -60m]	[170°, 190°]
			[80°, 100°]

TABLE I: Initial condition intervals

A. Simulation results

In the 100 simulations, the filter successfully tracked the target in all scenarios. Figure 5 and 6 show the actual and estimated target trajectory from a simulation with initial state $\boldsymbol{x}_0 = [-43.3\ 56.7\ 0.4\ -0.1\ -0.1\ 5.0]^\intercal$. A closer inspection of the trajectories, shown in Figure 6 reveals a zigzag pattern

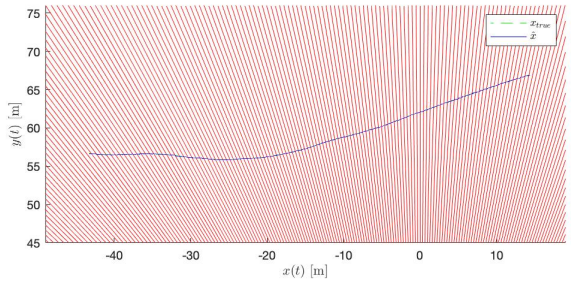


Fig. 5: Example of true and estimated trajectory.

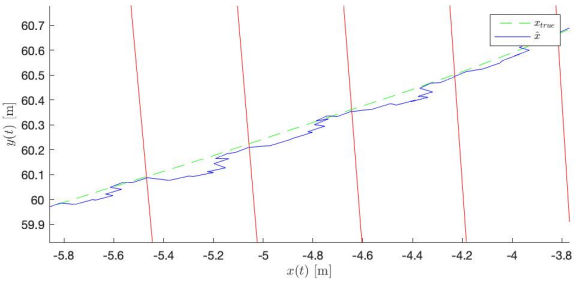


Fig. 6: Closer view of example of trajectories from Figure 5

that is typical for the all the estimated trajectories. The pattern arises from the assumption that the best guess for the predicted endpoint angles of the kayak is in the middle between two beams. While undesirable, it does not seem to negatively affect the overall performance of the filter.

The consistency of the filter was evaluated by calculating the averaged normalized estimation error squared (ANEES)

$$\text{ANEES}_k = \frac{1}{n_x T} \sum_{n=1}^{M} (\hat{x}_k^n - x_k^n)^\intercal \hat{\boldsymbol{P}}_k^{n\,-1} (\hat{x}_k^n - x_k^n). \tag{24}$$

where n_x is the number of states, T the number of time

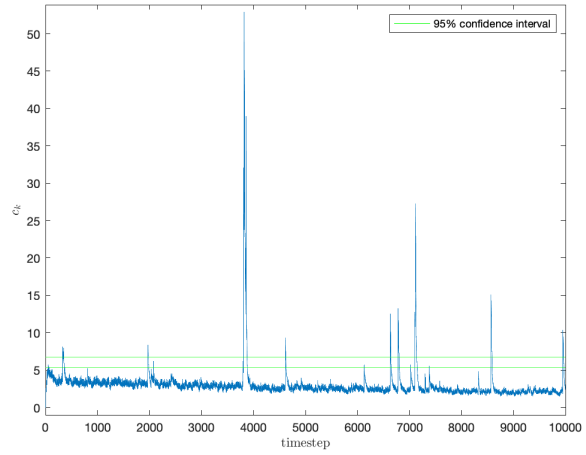


Fig. 7: Consistency analysis

steps and M the number of simulations. The resulting ANEES can be seen in Figure 7, where the 95% quantiles of the χ^2 distribution is shown in green. This indicate that the filter is somewhat underconfident. However it is clear from the figure that the filter does not diverge.

The root mean square error (RMSE) measure was also used to evaluate the filter's performance. For a state variable x the RMSE at time step k is given by

$$\text{RMSE}_k = \sqrt{\frac{1}{M}\sum_{n=1}^{M}(\hat{x}_k^n - x_k^n)^2}, \tag{25}$$

where M is total number of simulations. The time averaged RMSE for the different state variables are presented in Table II.

Variable	Value	Unit
x	0.0570	m
y	0.0577	m
u	0.1210	m/s
v	0.1160	m/s
ϕ	0.0925	rad
l	0.0022	m

TABLE II: Time averaged RMSE

V. Real Data test

The filter was also tested on real lidar data collected with a Velodyne VLP16 lidar with specifications as seen in Table III. The available data originated from an experiment where the lidar was positioned on a quay a small distance above the waterline, the setup and environment around the lidar is shown in Figure 8. The beams going parallel to the water was thus placed too high to detect the kayak. It was therefore decided to project the 3D measurements down to the 2D plane. Measurements from stationary objects was also filtered out before being sent to the filter.

Horizontal FOV	360°
Horizontal resolution	0.2°
Vertical FOV	+ 15.0° to -15.0°
Vertical resolution	2.0°

TABLE III: Specifications of lidar sensor used for data collection.

The results from running the filter on real data is shown in Figure 9. While the results seem promising, the sensor model used during simulations may be too simple as it does not include the possibility of the kayak being located between two of the vertical beam layers, i.e. that no beams will hit the kayak. This can be handled using for instance location dependent detection probability or more advanced sensor models while still preserving the essence of the model employed in this work.

VI. Conclusion and further work

This paper presented a direct approach for EOT using lidar measurements. The returns from each laser beam is processed directly in an EKF where data association is resolved using

Fig. 8: Lidar setup for data collection.

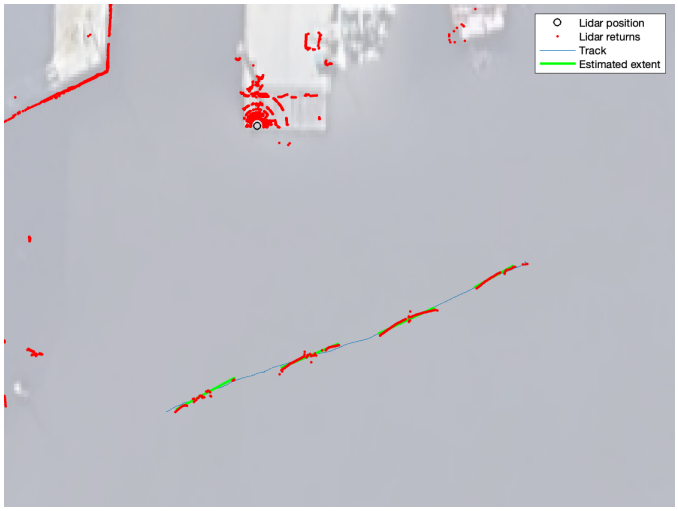


Fig. 9: Snapshot of estimated target and received measurements along with estimated target trajectory.

PDAF inspired techniques along with a RANSAC based method for hypothesis generation. We have shown that the suggested method is able to track a kayak through simulations and preliminary results for real data is also promising. Further work includes the evaluation of alternative methods such as the sample-based particle filter as a replacement for the EKF updates and the inclusion of other target models to account for non-stick shaped objects.

References

[1] K. Granström, M. Baum, and S. Reuter, "Extended Object Tracking: Introduction, Overview and Applications," *Journal of Advances in Information Fusion*, vol. 12, no. 2, pp. 139–174, 2017.

[2] K. Granström, S. Reuter, D. Meissner, and A. Scheel, "A multiple model PHD approach to tracking of cars under an assumed rectangular shape," in *17th International Conference on Information Fusion (FUSION)*. Salamanca: IEEE, 2014, pp. 1–8.

[3] K. Granström, S. Reuter, M. Fatemi, and L. Svensson, "Pedestrian tracking using Velodyne data - Stochastic optimization for extended

object tracking," in *Intelligent Vehicles Symposium (IV)*. Redondo Beach, CA: IEEE, 2017, pp. 39–46.

[4] K. Granström, L. Svensson, S. Reuter, Y. Xia, and M. Fatemi, "Likelihood-based data association for extended object tracking using sampling methods," *IEEE Transactions on Intelligent Vehicles*, vol. 3, no. 1, pp. 30 – 45, 2018.

[5] M. Schuster and J. Reuter, "Target tracking in marine environment using automotive radar and laser range sensor," in *20th International Conference on Methods and Models in Automation and Robotics (MMAR)*. Miedzyzdroje: IEEE, 2015, pp. 965–970.

[6] K. A. Ruud, E. F. Brekke, and J. Eidsvik, "LIDAR Extended Object Tracking of a Maritime Vessel Using an Ellipsoidal Contour Model," in *2018 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, no. 1. Bonn: IEEE, 2018.

[7] K. Granström, C. Lundquist, and U. Orguner, "Tracking Rectangular and Elliptical Extended Targets Using Laser Measurements," in *14th International Conference on Information Fusion*. Chicago, Illinois: IEEE, 2011, pp. 592–599.

[8] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*. Storrs, CT: YBS Publishing, 1995.

[9] E. Brekke and M. Chitre, "A multi-hypothesis solution to data association for the two-frame SLAM problem," *The International Journal of Robotics Research*, vol. 34, no. 1, pp. 43–63, 2015.

[10] A. Vedaldi, H. Tin, P. Favaro, and S. Soatto, "KALMANSAC: Robust filtering by consensus," in *Tenth IEEE International Conference on Computer Vision (ICCV'05)*, vol. 1. Beijing: IEEE, 2005, pp. 633–640.

[11] P. Niedfeldt and R. Beard, "Convergence and Complexity Analysis of Recursive-RANSAC: A New Multiple Target Tracking Algorithm," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 456 – 461, 2015.