Tracking Rectangular and Elliptical Extended Targets Using Laser Measurements

Karl Granström, Christian Lundquist, Umut Orguner
Department of Electrical Engineering, Linköping University, 581 33 Linköping, Sweden
Email: {karl,lundquist,umut}@isy.liu.se

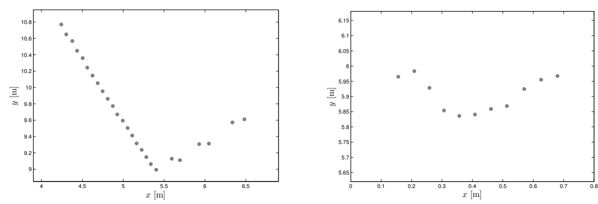
***Abstract*—This paper considers tracking of extended targets using data from laser range sensors. Two types of extended target shapes are considered, rectangular and elliptical, and a method to compute predicted measurements and corresponding innovation covariances is suggested. The proposed method can easily be integrated into any tracking framework that relies on the use of an extended Kalman filter. Here, it is used together with a recently proposed Gaussian mixture probability hypothesis density (GM-PHD) filter for extended target tracking, which enables estimation of not only position, orientation, and size of the extended targets, but also estimation of extended target type (i.e. rectangular or elliptical). In both simulations and experiments using laser data, the versatility of the proposed tracking framework is shown. In addition, a simple measure to evaluate the extended target tracking results is suggested.**

Keywords: Multiple target tracking, extended targets, probability hypothesis density, PHD, Gaussian mixture, Kalman filter, laser range data, rectangle, ellipse, intersection over union.

I. Introduction

Target tracking is the problem of estimating the states of an unknown number of targets using noisy and cluttered sets of measurements. In many typical target tracking scenarios the point target assumption is made, meaning that it is assumed that each target generates at most one measurement per time step. In recent years, tracking of extended targets have received increasing research attention. Here, extended target is defined as a target that potentially gives rise to more than one measurement per time step. Multiple measurements per target and time step enables the target tracking framework to not only estimate the location of each target, but also its spatial size and shape.

Gilholm and Salmond [1] presented an approach for tracking extended targets under the assumption that the number of recieved target measurements in each time step is Poisson distributed. They show an example where they track point targets which may generate more than one measurement, and an example where they track objects that have a 1-D extension (infinitely thin stick of length l). In [2] a measurement model was suggested which is an inhomogeneous Poisson point process. At each time step, a Poisson distributed random number of measurements are generated, distributed around the target. This measurement model can be understood to imply that the extended target is sufficiently far away from the sensor for its measurements to resemble a cluster of points, rather than a geometrically structured ensemble. A similar approach is taken in [3], where track-before-detect theory is used to track a point target with a 1-D extent.



(a) Measurements of a car. (b) Measurements of a human.

Figure 1. Motivating examples: measurements acquired outdoors with a laser range sensor. In (a) the measurements are approximately rectangular, and in (b) the measurements are approximately elliptical.

In [4] the authors model extended objects using an elliptic random hypersurface model [5]. Extended targets are modelled as having different measurement sources located on the target, and an ellipse is estimated that fits around the measurement sources. However, multiple extended target tracking is not treated. A Bayesian framework for estimating the location and radius of a circle from noisy measurements of the circle circumference is derived in [6]. A likelihood is not stated directly, instead the problem is posed using an errors-in-variables model.

In this work we consider estimation of extended targets using measurements from laser range sensors. Laser range sensor typically gives measurements with a high degree of structure, see examples in Figure 1, and are therefore suitable to use for the estimation of the shape and size of extended targets. In robotics laser range sensors have been used for tracking of vehicles and persons, we briefly present some more recent work here. The typical framework contains a detection algorithm that supplies a tracking algorithm with measurements belonging to some predefined class of targets. In [7], vehicles are modelled as rectangles and are tracked using a particle filter framework. In [8], work on detection and tracking of people and cars using a 2D laser range scanner and a camera is presented. However, there is no mention of estimation of the shape and size of the targets. Detection and tracking of pedestrians using 3D laser range data is presented in [9], where the position and velocity of the targets are tracked. Place dependent distributions of human behavior is used to improve tracking of people in [10]. The place dependency encodes locations in the surveillance area where the human targets are more likely to be located, and also captures areas where the human targets are less likely to move, e.g. crossing through walls.

Two types of extended targets are considered here, rectangular targets and elliptical targets. In computer vision tracking, rectangles and ellipses have been used as target bounding boxes, for targets of different types of shape. In comparison, here we are not concerned with estimation of target bounding boxes, but rather estimation of the shape, size and location of targets that are (approximately) rectangular or elliptical using point measurements. A similar scenario occurs in vision if feature points are extracted from the images, e.g. Harris corner points [11].

Here, the target type (ellipse or rectangle) is not detected from the measurements, but instead inferred in the target tracking estimation process. Bayesian estimation of extended targets from multiple measurements requires an appropriate likelihood function for the multiple measurements a target can generate. The paper presents a framework for computing these functions using the predicted measurements and corresponding innovation covariances for an extended target measured with a laser range sensor. It is also shown that this framework can successfully be integrated into an existing framework for extended target tracking, that is based on a Gaussian mixture Probability Hypothesis Density (GM-PHD) filter [12].

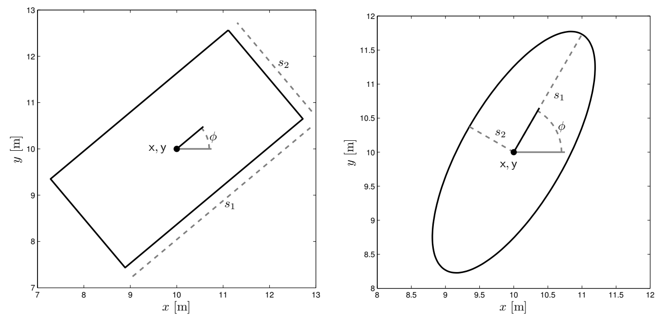
The extended target tracking framework is evaluated using both simulations and experiments. In simulations, single extended targets shaped as rectangles and ellipses are tracked. The results are evaluated against the ground truth. To evaluate the estimated shape and size of the extended targets, a performance metric called Intersection Over Union (IOU) is suggested. This measure is inspired by work in the computer vision research community, where it has been used to compare shapes to each other. In an outdoor experiment, up to three humans are tracked simultaneously, and the results are visually examined and shown to be good.

The paper is organized as follows: the next section defines the state representation of the extended targets, and relates the individual states in the state vector to the rectangular and elliptical shapes, respectively. Section III presents extended target tracking with a GM-PHD-filter, defines the main problem considered in this paper and addresses tracking of multiple extended targets with multiple shapes. In Section IV, a detailed implementational description is given of how predicted measurements and corresponding innovation covariances are computed. Section V presents a measure used for performance evaluation of extended target tracking, and Section VI presents results from both simulations and experiments. The paper is ended with conclusions and a discussion of future work in Section VII.

II. STATE REPRESENTATION

The state vector of a general two dimensional extended target is given as

$$\mathbf{x} = \begin{bmatrix} \mathsf{x} & \mathsf{y} & \mathsf{v}_\mathsf{x} & \mathsf{v}_\mathsf{y} & \phi & s_1 & s_2 \end{bmatrix}^{\mathrm{T}}, \tag{1}$$



(a) Rectangular target (b) Elliptical target

Figure 2. Relationship between target states (1) and target shape.

where x, y and $\mathsf{v}_\mathsf{x}, \mathsf{v}_\mathsf{y}$ are the Cartesian position and velocity in $\mathbb{R}^2$, respectively. The parameter ϕ is the orientation of the extended target shape, and s_1, s_2 represent the size of the shape. In this paper, two types of extended targets are considered, rectangular and elliptical shaped targets.

For rectangularly shaped extended targets, the two shape parameters s_1 and s_2 encode the length and width of the target, respectively, as is shown in Figure 2(a). Using this particular type of shape is motivated by the fact that cars measured by laser range sensors generate measurements that are approximately rectangularly shaped, see Figure 1(a).

For extended targets that are shaped like ellipsoids, the two shape parameters encode the lengths of the major and minor axis, respectively, as is shown in Figure 2(b). The ellipse shape for extended targets is motivated by the fact that humans measured by laser range sensors generate measurements that are approximately elliptically shaped, see Figure 1(b).

III. EXTENDED TARGET TRACKING

This section presents target tracking using GM-PHD-filters, defines the main problem addressed in the paper, and briefly considers estimation of multiple targets of multiple shapes.

A. GM-PHD *target tracking*

The aim is to estimate the state of a set of extended targets $\mathbf{X}_k = \left\{\mathbf{x}_k^{(j)}\right\}_{j=1}^{N_{x,k}}$ using sets of noisy, possibly cluttered, measurements $\mathbf{Z}_k = \left\{\mathbf{z}_k^{(j)}\right\}_{j=1}^{N_{z,k}}$, for discrete time instants $k = 1, \ldots, K$. In this paper, data from laser range sensors are used. Laser range sensors measure range r_i to the nearest object along rays pointing from the sensor at angles α_i. The measurements in $\mathbf{Z}_k$ can thus be sorted counter-clockwise according to the scanning angles α_i. Note that this order of the measurements does not contain information about which measurement source caused which measurement.

The target dynamics is assumed to be modelled with a function

$$\mathbf{x}_{k+1} = f\left(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k\right), \tag{2}$$

where $\mathbf{u}_k$ is an exogenous input and $\mathbf{w}_k$ is process noise with covariance matrix Q_k. State prediction using a dynamic motion model (2) is straightforward in target tracking, thus this part of the problem will not be adressed further in this publication.

In this work we have used the Gaussian Mixture Probability Hypothesis Density (GM-PHD) filter for extended target tracking presented in [12]. If $D_{k|k-1}(\mathbf{x}|\mathbf{Z})$ is the predicted PHD-intensity, the corrected PHD-intensity is

$$D_{k|k}(\mathbf{x}|\mathbf{Z}) = L_{\mathbf{Z}_k}(\mathbf{x})\, D_{k|k-1}(\mathbf{x}|\mathbf{Z}), \tag{3}$$

where the measurement pseudo-likelihood function [13] is given by

$$\begin{aligned} L_{\mathbf{Z}_k}(\mathbf{x}) =& 1 - \left(1 - e^{-\gamma(\mathbf{x})}\right) p_D(\mathbf{x}) + e^{-\gamma(\mathbf{x})} p_D(\mathbf{x}) \times \\ & \times \sum_{\mathrm{p}\angle\mathbf{Z}_k} \omega_{\mathrm{p}} \sum_{W\in\mathrm{p}} \frac{\gamma(\mathbf{x})^{|W|}}{d_W} \cdot \prod_{\mathbf{z}\in W} \frac{\phi_{\mathbf{z}}(\mathbf{x})}{\lambda_k c_k(\mathbf{z})}. \end{aligned} \tag{4}$$

The first part of this equation, $1-\left(1-e^{-\gamma(\mathbf{x})}\right)p_D(\mathbf{x})$, handles the targets for which there are no detections. The second part handles targets for which there are at least one detection. The PHD-intensity is approximated by a Gaussian mixture as in [14],

$$D_k(\mathbf{x}) = \sum_{i=1}^{J_k} w_k^{(i)} \mathcal{N}\left(\mathbf{x}\,;\, m_k^{(i)}, P_k^{(i)}\right), \tag{5}$$

where $w_k^{(i)}$, $m_k^{(i)}$ and $P_k^{(i)}$ are the weights, mean vectors and covariance matrices of the Gaussian components, respectively.

As (4) shows, the update step contains a summation over partitions p of the measurement set $\mathbf{Z}_k$, and a summation over the cells W in each partition p. Let the measurements in a cell W be denoted

$$\mathbf{z}^W \triangleq \bigoplus_{\mathbf{z}\in W} \mathbf{z}, \tag{6}$$

where $\bigoplus$ is vertical vectorial concatenation. Then, for each cell of each partition and each predicted Gaussian component,

$$\begin{aligned} & e^{-\gamma(\mathbf{x})} p_D(\mathbf{x})\, \omega_{\mathrm{p}} \frac{\gamma(\mathbf{x})^{|W|}}{d_W} \times \\ & \times \frac{\mathcal{N}\left(\mathbf{z}^W\,;\, \hat{\mathbf{z}}_{k|k-1}^{W,(i)}, S_k^{W,(i)}\right)}{(\lambda_k c_k(\mathbf{z}))^{|W|}} \mathcal{N}\left(\mathbf{x}\,;\, \hat{\mathbf{x}}_{k|k}^{(i)}, P_{k|k}^{(i)}\right) \end{aligned} \tag{7}$$

represents the corresponding updated Gaussian component, where $\mathcal{N}\left(\mathbf{z}^W\,;\, \hat{\mathbf{z}}_{k|k-1}^{i}, S_k^{W,(i)}\right)$ is the likelihood of the set of measurements in the cell W. The key point of being able to use the above formulas for extended targets with structured measurements (e.g., with laser sensor reports) is to calculate the predicted measurements $\hat{\mathbf{z}}_{k|k-1}^{W,(i)}$, innovation covariances $S_k^{W,(i)}$, updated means $\hat{\mathbf{x}}_{k|k}^{(j)}$ and updated covariances $P_{k|k}^{(i)}$. The calculation of these quantities must rely on a measurement model of a form similar to

$$\mathbf{z}_k^{(j)} = h(\mathbf{x}_k, \mathbf{e}_k), \tag{8}$$

where $\mathbf{e}_k$ is measurement noise with covariance matrix $R_k^{(j)}$. Once such a model is available,

- the calculation of the measurement prediction $\hat{\mathbf{z}}_{k|k-1}^{W,(i)}$ and innovation covariance $S_k^{W,(i)}$ can be achieved using the current estimates, and
- as in the implementation presented in [12], the updated means $\hat{\mathbf{x}}_{k|k}^{(i)}$ and updated covariances $P_{k|k}^{(i)}$ can be obtained with a Kalman filter, or one of its non-linear counterparts such as the Extended Kalman Filter (EKF) or Unscented Kalman Filter (UKF).

The construction of a measurement model of the form (8) is at the root of the problem of calculating a measurement prediction which will be posed as the main problem of our work in the next subsection.

B. Problem definition

When the extended targets are modeled as points, a measurement model can readily be constructed assuming that each measurement is close to the target's center of mass, see e.g. [12]. Here however, we are concerned with tracking extended targets (1) using measurements acquired with a laser range sensor, thus the point target assumption is trivially invalid. As will be shown using an example, when the extended targets are not modeled as points but rather as geometric structures, construction of a measurement model can be significantly more complicated compared to the point target case.

Consider the plots in Figure 3. The extended target estimate is shifted in y position by $0.7\mathrm{m}$, and in heading ϕ by $20°$, with respect to the true extended target. Using a simple model of a laser range sensor, predicted measurements can be computed using line intersection as is shown in Figure 3(a). However, due to the orientation error, the estimated target is not showing the same two sides towards the sensor as the true target, making the set of true measurements $\mathbf{Z}_k$ fundamentally different from the set of estimated measurements $\hat{\mathbf{Z}}_k$. The cardinality is different, $|\mathbf{Z}_k| = 16$ and $|\hat{\mathbf{Z}}_k| = 14$, and the measurements are located on different sides of the target shape, causing the data association between the measurements in $\mathbf{Z}_k$ and the predicted measurements in $\hat{\mathbf{Z}}_k$ to become difficult. However, using the method proposed in this paper, described further in Section IV, a $\hat{\mathbf{Z}}_k$ that better corresponds to $\mathbf{Z}_k$ is obtained, see Figure 3(b).

The main problem addressed in this paper is thus to obtain predicted measurements and innovation covariances such that they can be integrated into an existing target tracking framework. As mentioned above, here the GM-PHD extended target tracking framework presented in [12] has been used, however any target tracking framework that relies on the EKF could have been used just as well. In order to obtain predicted measurements $\hat{\mathbf{z}}_{k|k-1}^{W,(i)}$ and corresponding innovation covariances $S_k^{W,(i)}$, information from the set of measurements in the cell $\mathbf{z}^W$ is used, i.e. the following two approximations

$$\hat{\mathbf{z}}_{k|k-1}^{W,(i)} \approx \hat{\mathbf{z}}_{k|k-1}^{W,(i)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}\right) \tag{9a}$$

$$S_k^{W,(i)} \approx S_k^{W,(i)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}\right) \tag{9b}$$

are made. Thus, the predicted measurements and innovation covariances are functions of the predicted state and the set

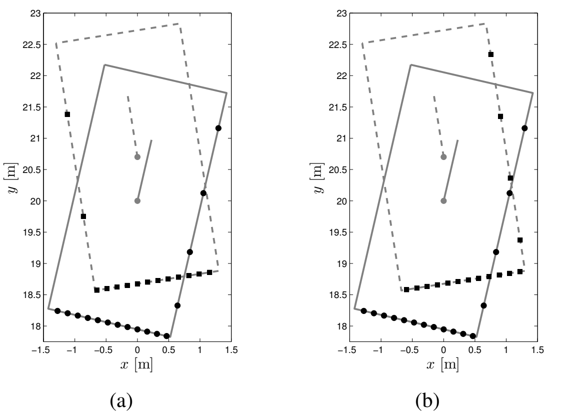


Figure 3. Comparison of estimated measurements, the sensor is located in the origin. The true extended target is shown as a solid rectangle and the set of measurements $\mathbf{Z}_k$ are shown as filled black circles. The estimated extended target is shown as a dashed rectangle, the estimated set of measurements $\hat{\mathbf{Z}}_k$ are shown as filled black squares. In (a), using a straightforward model of the laser sensor gives a $\hat{\mathbf{Z}}_k$ which is a poor correspondence to $\mathbf{Z}_k$. In (b), the proposed method is used and $\hat{\mathbf{Z}}_k$ better corresponds to $\mathbf{Z}_k$.

of measurements. In a sense, this is an unconventional use of an errors-in-variables framework where the measurement model would depend on the current set of measurements. In Section IV we describe in detail the proposed method for computing $\hat{\mathbf{z}}_{k|k-1}^{W,(i)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}\right)$ and $S_k^{W,(i)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}\right)$. When these quantities are calculated, a Kalman filter (or EKF/UKF) measurement update can easily be used to obtain the updated means $\hat{\mathbf{x}}_{k|k}^{(i)}$ and updated covariances $P_{k|k}^{(i)}$.

C. Multiple shapes and multiple targets

An interesting aspect of extended target tracking with multiple targets and multiple shapes is how to correctly estimate the correct shape for each target. As was mentioned in the introduction, in some previous work the target tracking is preceeded by a detection algorithm. Thus, a possible way to infer the type of shape is to consider the measurements and make a hard decision as to which type they represent.

In this paper, an approach that is slightly similar to track-before-detect is taken. When a new target appears, the GM-PHD-filter birth intensity is set such that one Gaussian component per target type is given birth to. Then, as the filter iterates throughout the prediction and correction steps, the weight $w_k^{(j)}$ of each Gaussian component is updated. Eventually, the weights will converge such that only one component remains, from which the target type can be found. The target types are thus, while not included in the state vectors, implicitly estimated via the weights of the Gaussian components in the PHD-intensity.

IV. Computing predicted measurements and innovation covariances

In this section we present a detailed description of how predicted measurements and innovation covariances (9) are computed for extended targets (1) that are either rectangular or elliptical in shape.

A. Predicted measurements for rectangular targets

The first step in predicting a set of measurements is to, given the set of measurements, estimate how many sides of the measured target that are shown. For a rectangular target, it is trivial to conclude that at most two of the sides of the shape are visible to the sensor at any given moment. Given a set of measurements $\mathbf{Z} = \left\{\mathbf{z}_k^{(j)}\right\}_{j=1}^{N_{z,k}}$, where each measurement is a vector $\mathbf{z}_k^{(j)} = \left[z_k^{(j),1}, \ldots, z_k^{(j),n_z}\right]^{\mathrm{T}}$, let $C = [c_{m,n}]$ be the sample covariance of the measurements with entries

$$c_{m,n} = \frac{1}{N_{z,k}-1}\sum_{j=1}^{N_{z,k}}\left(z_k^{(j),m} - \bar{z}_k^m\right)\left(z_k^{(j),n} - \bar{z}_k^n\right) \quad (10)$$

where $\bar{z}_k^m$ is the mean of the m:th component of the measurement vectors $\mathbf{z}_k^{(j)}$. For laser range measurements in 2D we have $n_z = 2$.

Further, let e_1 and e_2 be the two eigenvalues of the covariance matrix C, where $e_2 > e_1$. In the noiseless case, measurements of just one side of a rectangle will have 0 standard deviation along the direction perpendicular to the measured line, and the corresponding eigenvalue will be 0. An estimate of the number of sides N that generated the set of measurements can then be obtained as

$$N = \begin{cases} 1 & \text{if } \frac{e_2}{e_1} \geq K \\ 2 & \text{otherwise} \end{cases} \quad (11)$$

where K is a threshold. Empirically $K = 25$ was determined to be an appropriate value. Given the measurements in Figure 4(a)[1], the eigenvalues of the corresponding sample covariance matrix are $e_1 = 0.1198$ and $e_2 = 0.6629$, and the eigenvalue quota 5.5327 is less than the threshold K.

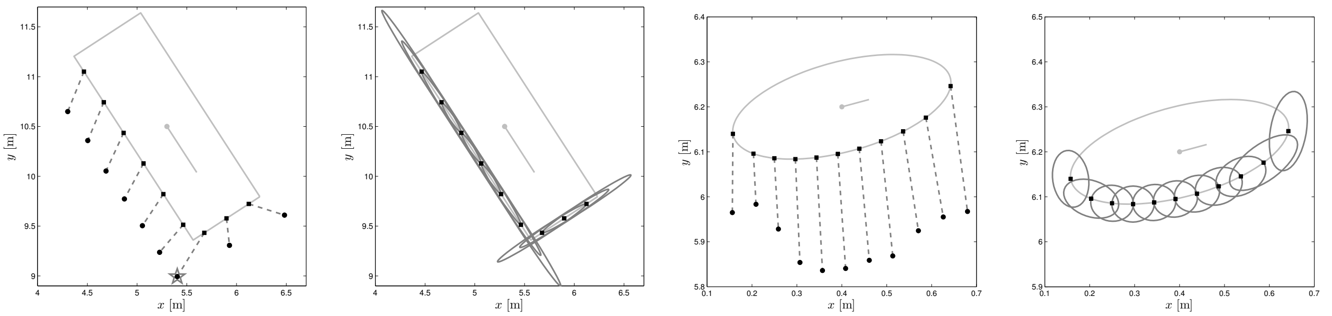
If the estimated number of sides is two, the measurement closest to the corner connecting the two sides needs to be found. This is performed using the function CORNER INDEX given in Table I, where the distance d from a point $\mathbf{z}_3$ to a line defined by two points $\mathbf{z}_1$ and $\mathbf{z}_2$ is given by the function

$$\begin{aligned} d &= \text{point2lineDist}\left(\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\right) \\ &= \frac{\left|\left(\mathbf{z}_2^{\mathsf{x}} - \mathbf{z}_1^{\mathsf{x}}\right)\left(\mathbf{z}_1^{\mathsf{y}} - \mathbf{z}_3^{\mathsf{y}}\right) - \left(\mathbf{z}_1^{\mathsf{x}} - \mathbf{z}_3^{\mathsf{x}}\right)\left(\mathbf{z}_2^{\mathsf{y}} - \mathbf{z}_1^{\mathsf{y}}\right)\right|}{\sqrt{\left(\mathbf{z}_2^{\mathsf{x}} - \mathbf{z}_1^{\mathsf{x}}\right)^2 + \left(\mathbf{z}_2^{\mathsf{y}} - \mathbf{z}_1^{\mathsf{y}}\right)^2}}. \end{aligned} \quad (12)$$

For the measurements in Figure 4(a), the corner is identified as the point located in $x = 5.40, y = 8.99$. Given the identified corner, it is straightforward to compute the number of measurements belonging to each of the two sides that are seen by the sensor. In the example given in Figure 4(a), there are $m_1 = 3$ measurements on one side, and $m_2 = 6$ on the other. If only one side is measured, i.e. if $\frac{e_2}{e_1} \geq K$, then trivially all measurements belong to the side that was measured by the sensor.

Assuming that two sides are shown by the set of measurements $\mathbf{Z}$, the set can be divided into two subsets $\mathbf{Z}^1$ and $\mathbf{Z}^2$

[1]The measurements are a subset of the measurements in Figure 1(a). A subset is used to prevent Figures 4(a) and 4(b) from being too cluttered.



(a) $\hat{\mathbf{Z}}$ (squares) and $\mathbf{Z}$ (circles) (b) Measurement covariances (c) $\hat{\mathbf{Z}}$ (squares) and $\mathbf{Z}$ (circles) (d) Measurement covariances

Figure 4. The rectangular and elliptical targets. (a) and (c): Example of measurements of a car and human (filled black circles), respectively, and predicted measurements located on target surface (filled black squares). The corresponding associations are shown with dashed lines. The measurement identified as being closest to the corner in the rectangular case is shown with a gray star. (b) and (d): The corresponding measurement covariances.

Table I
CORNER INDEX

Input: Set of measurements $\mathbf{Z} = \{\mathbf{z}^{(j)}\}_{j=1}^{N_z}$, sorted counter clockwise according to scanning angle.
Initialise: Minimum distance $d_{\text{min}} = \infty$, first point $\mathbf{z}_1 = \mathbf{z}^{(1)}$ and last point $\mathbf{z}_{N_z} = \mathbf{z}^{(N_z)}$.

for $n = 2, \ldots, N_z - 1$ **do**
  Let $\mathbf{z}_n = \mathbf{z}^{(n)}$ be the current point.
  Initialise sum of distances $d = 0$.
  for $k = 2, \ldots, n-1, n+1, \ldots, N_Z - 1$ **do**
    $d = d + \begin{cases} \text{point2lineDist}\,(\mathbf{z}_1, \mathbf{z}_n, \mathbf{z}_k) & \text{if } k < n \\ \text{point2lineDist}\,(\mathbf{z}_n, \mathbf{z}_{N_z}, \mathbf{z}_k) & \text{if } k > n \end{cases}$
  end for
  if $d < d_{\text{min}}$ **then**
    Set $\hat{n} = n$ and $d_{\text{min}} = d$
  end if
end for

Output: Index to measurement closest to corner $\hat{n}$.

corresponding to the measurements that are from the two sides, i.e. $|\mathbf{Z}^1| = m_1$ and $|\mathbf{Z}^2| = m_2$. Let β_1 and β_2 be the angles of the vectors defined by the first and last point from $\mathbf{Z}^1$ and $\mathbf{Z}^2$, respectively. Then, $\beta_i^N = \beta_i + \pi/2$ are the angles of the corresponding normal vectors. Further, given an estimated extended target state $\hat{\mathbf{x}}$, let $\xi_1, \ldots, \xi_4$ be the surface normals of the four sides of the rectangle.

The sets $\mathbf{Z}^1$ and $\mathbf{Z}^2$ can now be associated to one of the four sides of the rectangle by finding the two sides for which $\left|\xi_i - \beta_1^N\right|$ and $\left|\xi_j - \beta_2^N\right|$ are minimised. Predicted measurements are generated for the sides that are in view by distributing m_1 and m_2 points uniformly on the two sides. An example is given in Figure 4(a).

B. Predicted measurements for elliptical targets

Given an angle at which the sensor measures, finding the intersection between the measurement ray and an ellipse is performed as follows. Let the ellipse be given by the position x, y, orientation ϕ and lengths of the major and minor axis s_1 and s_2, as in (1). Further, let the range measurement r from the sensor to the target surface be measured at an angle α. Thus, the intersection defines a point in Cartesian coordinates

$$\begin{bmatrix} \mathsf{x}_r \\ \mathsf{y}_r \end{bmatrix} = \begin{bmatrix} r\cos(\alpha) \\ r\sin(\alpha) \end{bmatrix}, \tag{13}$$

given in the coordinate frame defined with the sensor position as origin. The same point can be described in the coordinate frame defined by the position and heading of the ellipse, using the appropriate coordinate frame transformation. The intersection point's coordinates are now given by

$$\begin{aligned} \begin{bmatrix} \mathsf{x}_r^e \\ \mathsf{y}_r^e \end{bmatrix} &= (\mathsf{R}_{-\alpha})^{-1} \begin{bmatrix} \mathsf{x}_r - \mathsf{x} \\ \mathsf{y}_r - \mathsf{y} \end{bmatrix} \\ &= \begin{bmatrix} r\,(c_\alpha c_\phi + s_\alpha s_\phi) - \mathsf{x} c_\phi - \mathsf{y} s_\phi \\ r\,(-c_\alpha s_\phi + s_\alpha c_\phi) + \mathsf{x} s_\phi - \mathsf{y} c_\phi \end{bmatrix} = \begin{bmatrix} r\theta_1 + \theta_2 \\ r\theta_3 + \theta_4 \end{bmatrix} \end{aligned} \tag{14}$$

where R_α is the rotation matrix for an angle α and e denotes the change of reference frame. This point must satisfy the ellipse equation

$$\frac{(\mathsf{x}_r^e)^2}{s_1^2} + \frac{(\mathsf{y}_r^e)^2}{s_2^2} = 1. \tag{15}$$

Inserting (14) into (15) gives

$$\begin{aligned} 1 &= \frac{(r\theta_1 + \theta_2)^2}{s_1^2} + \frac{(r\theta_3 + \theta_4)^2}{s_2^2} \\ &= \left(\frac{\theta_1^2}{s_1^2} + \frac{\theta_3^2}{s_2^2}\right) r^2 + 2\left(\frac{\theta_1\theta_2}{s_1^2} + \frac{\theta_3\theta_4}{s_2^2}\right) r + \frac{\theta_2^2}{s_1^2} + \frac{\theta_4^2}{s_2^2} \\ &= Ar^2 + 2Br + C \end{aligned} \tag{16}$$

which has the two solutions $r = -\frac{B}{A} \pm \sqrt{\frac{B^2}{A^2} - \frac{C-1}{A}}$. Since the sensor measures the closest intersection with the target, the correct range r at a given angle α is $r = -\frac{B}{A} - \sqrt{\frac{B^2}{A^2} - \frac{C-1}{A}}$. Note that if no part of the extended target is located along the measurement ray defined by the angle α, r will be a complex number.

For elliptically shaped extended targets, the first step in computing a set of predicted measurements is to find the angles α_1 and α_2 within which the sensor can measure the estimated target. Next, $|\mathbf{Z}|$ estimated measurements are generated on the estimated target surface uniformly spaced in the angle dimension between α_1 and α_2. An example of predicted measurements is given in Figure 4(c).

C. Innovation covariances

To compute the innovation covariances $S_k^{W,(j)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}\right)$, the measurement model Jacobian H_k and the measurement covariances $R_k^{(j)}$ are needed.

As we have not derived an explicit mathematical function for the measurement model, the measurement model Jacobian

$$H_k = \left.\frac{dh}{d\mathbf{x}}\right|_{\hat{\mathbf{x}}_{k|k-1}}, \tag{17}$$

is computed numerically instead of derived analytically. By making a small permutation ε to the n:th element of the predicted state vector, the n:th column of the measurement Jacobian is approximated as

$$\frac{\hat{\mathbf{z}}_{k|k-1}^{W,(j)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}^{\varepsilon,n}\right) - \hat{\mathbf{z}}_{k|k-1}^{W,(j)}\left(\mathbf{z}^W, \hat{\mathbf{x}}_{k|k-1}\right)}{\varepsilon} \tag{18}$$

where $\hat{\mathbf{x}}_{k|k-1}^{\varepsilon,n}$ is the predicted state after the n:th element is permuted by adding ε. Note that in doing so, we only consider permutations in the predicted state when the Jacobian is approximated numerically, i.e. we do not consider permutations in the set of measurements $\mathbf{z}^W$. Considering changes in $\mathbf{z}^W$ is a topic for future work.

Instead of modelling the covariances according to the sensor statistics, the covariances are constructed such that the uncertainty of all predicted measurements follow the surface of the target. One of the axes of the corresponding uncertainty ellipse is aligned with the surface, and the size of the uncertainty in this direction is set to the distance to the nearest measurement point. The size of the uncertainty in the direction perpendicular to the surface is set to a constant σ_r.

With multiple measurements on the target surface, this gives measurement covariance ellipses that are aligned to the surface tangent, thus giving a combined uncertainty that covers the part of the target surface that was measured by the sensor. Examples of measurement covariances for the measurements given in Figures 4(a) and 4(c) are shown in Figures 4(b) and 4(d).

V. Extended target tracking performance evaluation

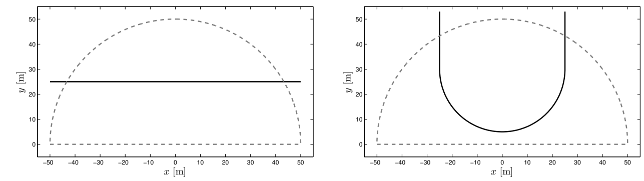
In this section we address evaluating the estimated shape of the extended target. As the velocity of the extended target does not affect its shape and size, the two velocity states v_x and v_y are excluded here for the sake of simplicity. For a rectangular or elliptical extended target with true state $\mathbf{x}^0 = \begin{bmatrix}\mathsf{x}^0 & \mathsf{y}^0 & \phi^0 & s_1^0 & s_2^0\end{bmatrix}^{\mathrm{T}}$, there are four alternative estimates that give an identical shape in the state space:

$$\hat{\mathbf{x}}^1 = \begin{bmatrix}\mathsf{x}^0 & \mathsf{y}^0 & \phi^0 & s_1^0 & s_2^0\end{bmatrix}^{\mathrm{T}}, \tag{19a}$$

$$\hat{\mathbf{x}}^2 = \begin{bmatrix}\mathsf{x}^0 & \mathsf{y}^0 & \phi^0 + \pi & s_1^0 & s_2^0\end{bmatrix}^{\mathrm{T}}, \tag{19b}$$

$$\hat{\mathbf{x}}^3 = \begin{bmatrix}\mathsf{x}^0 & \mathsf{y}^0 & \phi^0 + \frac{\pi}{2} & s_2^0 & s_1^0\end{bmatrix}^{\mathrm{T}}, \tag{19c}$$

$$\hat{\mathbf{x}}^4 = \begin{bmatrix}\mathsf{x}^0 & \mathsf{y}^0 & \phi^0 - \frac{\pi}{2} & s_2^0 & s_1^0\end{bmatrix}^{\mathrm{T}}. \tag{19d}$$



(a) Linear motion (b) Linear and curved motion

Figure 6. Trajectories used in simulations. The target trajectories are showed in black. The sensor is located in the origin, the surveilance area boundary is showed with a dashed gray line.

For example, let the orientation of a rectangular target be $\phi = 0\mathrm{rad}$ and let the lengths of the two sides be $s_1 = 4\mathrm{m}$ and $s_2 = 2\mathrm{m}$. The estimation errors of the corresponding estimated states may be as large as $\frac{\pi}{2}\mathrm{rad}$, $-2\mathrm{m}$ and $2\mathrm{m}$, despite the fact that if visualised, the estimated shape and size of the extended target is identical to the true one. Thus, only considering the estimation errors of the extended target tracking results may give a false picture of the quality of the extended target tracking results.

In this paper, we evaluate the extended target tracking results by considering the estimated x, y-position, and by considering a measure called Intersection-Over-Union (IOU). Let $\hat{A}$ be the area of the estimated extended target, and let A_0 be the area of the true target. By computing the area of the intersection between the estimate and the true target, and dividing by the union of the two areas,

$$\frac{\hat{A} \cap A_0}{\hat{A} \cup A_0} \in [0\ 1] \tag{20}$$

a measure is obtained, where 1 represents a perfect overlap of the estimated and true extended target, and 0 represents that there is no overlap at all. The IOU-measure captures differences in x, y-position, in shape orientation ϕ and in the shape size parameters s_1 and s_2. It should be noted though that the IOU-measure does not consider any difference in target type, thus the measure could very well be close to one despite the fact that the true target is rectangular and the estimated target is elliptical, or vice versa.

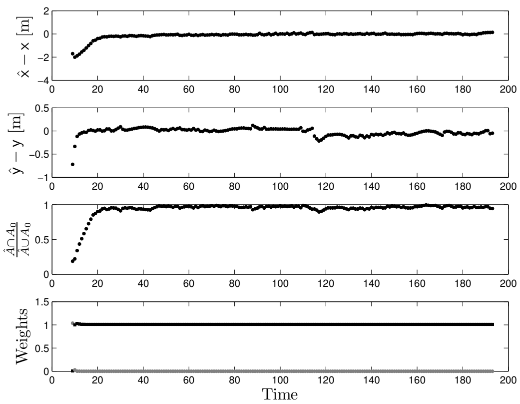
The estimated target type is evaluate by considering the sum of the Gaussian components' weights $w_k^{(j)}$ for each type. Assuming that there is only one rectangular target present, the weights for the Gaussian components representing rectangular targets should sum to one, and the weights for the components representing elliptical targets should sum to zero.

VI. Results

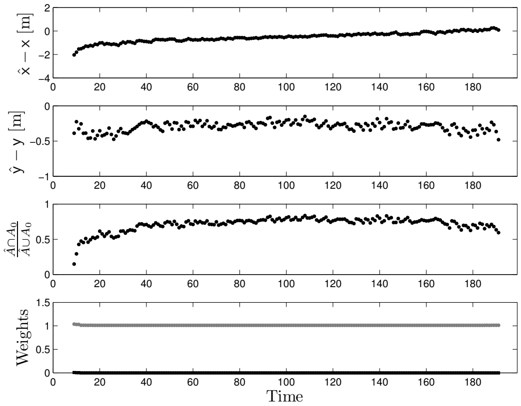
This section presents results from simulations and experiments using the presented work.

A. Simulations

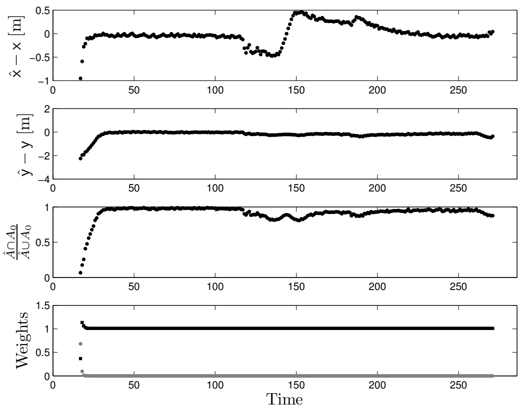
A number of different simulations were performed in order to assess the estimation results of the extended target tracking filter. Two simulated trajectories were used, one linear motion, shown in Figure 6(a), and a combination of linear and curved motion, shown in Figure 6(b).



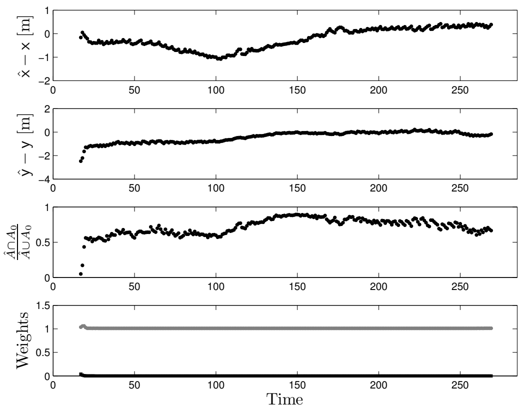
(a) Rectangular target, linear motion



(b) Elliptical target, linear motion



(c) Rectangular target, linear and curved motion



(d) Elliptical target, linear and curved motion

Figure 5. Simulation results: the top row shows linear motion (Figure 6(a)), the bottom row linear and curved motion (Figure 6(b)). Each figure shows the estimation error in x and y position, the IOU measure and the sum of the weights for each target type (black is rectangles, gray is ellipses).

1) Linear motion: In the first simulation a rectangularly shaped extended target moves from right to left throught the surveilance area. The orientation of the target is $0\mathrm{rad}$ and the length and width is $s_1 = 5\mathrm{m}$ and $s_2 = 2.5\mathrm{m}$, respectively. The estimation results are shown in Figure 5(a). A similar simulation was performed using an elliptical extended target. The motion is again from right to left, and the orientation is $0\mathrm{rad}$. The lengths of the major and minor axes are $s_1 = 2.5\mathrm{m}$ and $s_2 = 1.25\mathrm{m}$, respectively. The estimation results are shown in Figure 5(b).

2) Linear and curved motion: In this simulation the motion of the extended target was a combination of linear and curved motion. A rectangular target with length and width $s_1 = 5\mathrm{m}$ and $s_2 = 2.5\mathrm{m}$ was simulated first, the estimation results are shown in Figure 5(c). In a similar simulation, an elliptical target with lengths of the major and minor axes $s_1 = 2.5\mathrm{m}$ and $s_2 = 1.25\mathrm{m}$, respectively, was simulated. The estimation results are shown in Figure 5(d).

3) Comments: Under linear motion, the rectangular target is estimated with high accuracy, while the elliptical target is slightly underestimated in size (s_1 and s_2 are underestimated). Under linear and curvet motion the problem is slightly more complicated for the rectangular target as it moves close to the sensor, however the overall results are good. For the elliptical target, the size is underestimated again. Estimation of target type (rectangular or elliptical) is shown via the Gaussian component weights in the bottom plots in each figure. As is shown, the filter quickly converges to the correct target type.

B. Experiment

The suggested framework for extended target tracking was tested in an experiment using laser range data. The data set used contains 600 range scans acquired in an outdoor environment, with five persons moving through the surveilance area (at most three persons simultaneously). The first person enters the surveilance area at time 22 and moves to the center where he remains still for the remainder of the data. The second person enters at time 38, and proceeds to move behind the first person, both entering and exiting an occluded part of the surveilance area. Remaining three persons enters and exits the surveilance area at later times during the experiment. The results from the experiment are shown in Figures 7 and 8.

Since the second person moves through parts of the surveilance area that are occluded by the first person, the results show target loss (i.e. cardinality error) at three time steps. Using a variable probability of detection, this problem

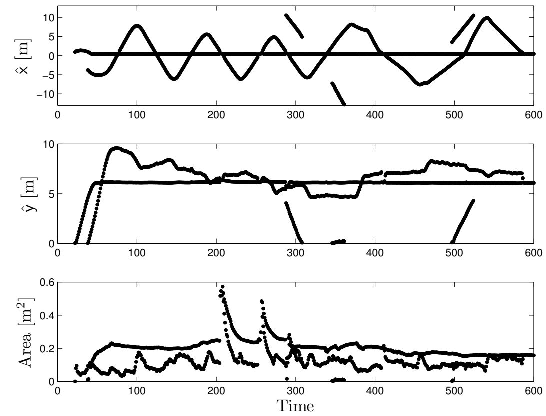


Figure 7. Experiment results, showing the estimated x and y position, as well as the area of each extended target.

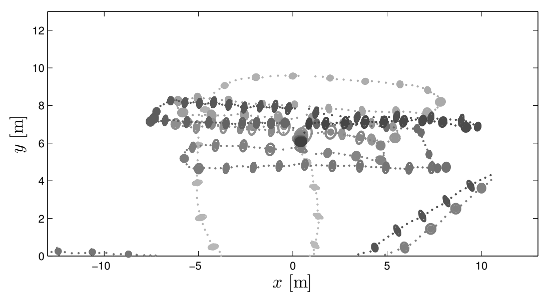


Figure 8. Experiment results, showing the trajectories of each of the two targets. A grayscale is used to highlight the different time steps, and for every 5:th time step the corresponding shape and size of each target is plotted.

can be overcome. However we have chosen to not include this, since the space constraints do not allow a description of the method used to compute the variable probability of detection.

As there is no ground truth, it is difficult to evaluate the quality of the estimated extended target, including the shape parameters ϕ, s_1 and s_2. In this paper we have chosen to compute the area of each extended target, since this area can be compared to a rough estimate of the area of a cross section of the human torso, under the assumption that it is elliptically shaped. Under the assumption that an average person is roughly $50\mathrm{cm}$ to $60\mathrm{cm}$ wide (torso and arms) and $25\mathrm{cm}$ to $30\mathrm{cm}$ deep, the average area can be said to be somewhere between $0.1\mathrm{m}^2$ and $0.15\mathrm{m}^2$. Comparing these values to the estimated areas (see Figure 7) show that the ellipses representing the persons have estimated areas of resonable size, with the exception of two instances in time between time 200 and time 300. These two times correspond to time when two targets are spatially very close, and are thus merged into just one target, producing a considerably larger target. The person standing still shows a stable estimated area, the persons that are moving throughout the whole experiment shows much more changes in the estimated area.

VII. Conclusions and Future Work

In this paper we presented a method to compute predicted measurements and corresponding innovation covariances when rectangular and elliptical extended targets are measured by laser range sensors. The method can easily be inserted to an existing extended target tracking GM-PHD-filter, enabling efficient estimation of the extended target's location, orientation and size, as well as estimation of the extended target type.

In future work, we plan to investigate the reasons behind the underestimation of target size for elliptical targets. Furthermore, the presented work needs to be integrated with the variable probability of detection such that targets can be tracked while they are occluded by other targets. The target tracking framework also needs to be tested in experiments with laser range data that contains measurements of both rectangular and elliptical targets, in order to test the estimation of target type further.

Acknowledgement

This work has been supported by the Swedish Research Council under the Linnaeus Center CADICS.

References

[1] K. Gilholm and D. Salmond, "Spatial distribution model for tracking extended objects," *IEE Proceedings Radar, Sonar and Navigation*, vol. 152, no. 5, pp. 364–371, Oct. 2005.

[2] K. Gilholm, S. Godsill, S. Maskell, and D. Salmond, "Poisson models for extended target and group tracking," in *Proceedings of Signal and Data Processing of Small Targets*, vol. 5913. San Diego, CA, USA: SPIE, Aug. 2005, pp. 230–241.

[3] Y. Boers, H. Driessen, J. Torstensson, M. Trieb, R. Karlsson, and F. Gustafsson, "A track before detect algorithm for tracking extended targets," *IEE Proceedings Radar, Sonar and Navigation*, vol. 153, no. 4, pp. 345–351, Aug. 2006.

[4] M. Baum, B. Noack, and U. D. Hanebeck, "Extended Object and Group Tracking with Elliptic Random Hypersurface Models," in *Proceedings of the International Conference on Information Fusion*, Edinburgh, UK, Jul. 2010.

[5] M. Baum and U. D. Hanebeck, "Random hypersurface models for extended object tracking," in *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Ajman, United Arab Emirates, Dec. 2009, pp. 178–183.

[6] M. Baum, V. Klumpp, and U. D. Hanebeck, "A Novel Bayesian Method for Fitting a Circle to Noisy Points," in *Proceedings of the International Conference on Information Fusion*, Edinburgh, UK, Jul. 2010.

[7] A. Petrovskaya and S. Thrun, "Model Based Vehicle Tracking for Autonomous Driving in Urban Environments," in *Proceedings of Robotics: Science and Systems*, Zürich, Switzerland, Jun. 2008.

[8] L. Spinello, R. Triebel, and R. Siegwart, "Multiclass Multimodal Detection and Tracking in Urban Environments," *International Journal for Robotics Research*, vol. 29, no. 12, pp. 1498–1515, Oct. 2010.

[9] L. E. Navarro-Serment, C. Mertz, and M. Hebert, "Pedestrian Detection and Tracking Using Three-dimensional LADAR Data," *International Journal for Robotics Research*, vol. 29, no. 12, pp. 1516–1528, Oct. 2010.

[10] M. Luber, G. D. Tipaldi, and K. O. Arras, "Place-dependent people tracking," *International Journal for Robotics Research*, vol. 30, no. 3, pp. 280–293, Jan. 2011.

[11] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of Alvey Vision Conference*, Manchester, UK, Sep. 1988, pp. 147–152.

[12] K. Granström, C. Lundquist, and U. Orguner, "A Gaussian mixture PHD filter for extended target tracking," in *Proceedings of the International Conference on Information Fusion*, Edinburgh, UK, Jul. 2010.

[13] R. Mahler, "PHD filters for nonstandard targets, I: Extended targets," in *Proceedings of the International Conference on Information Fusion*, Seattle, WA, USA, Jul. 2009, pp. 915–921.

[14] B.-N. Vo and W.-K. Ma, "The Gaussian mixture probability hypothesis density filter," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4091–4104, Nov. 2006.