

Tracking and Data Segmentation Using a GGIW Filter with Mixture Clustering

Alexander Scheel*, Karl Granström†, Daniel Meissner*, Stephan Reuter*, Klaus Dietmayer*

*Institute of Measurement, Control, and Microtechnology

Ulm University, Ulm, Germany

Email: {alexander.scheel, daniel.meissner, stephan.reuter, klaus.dietmayer}@uni-ulm.de

†Department of Electrical Engineering

Linköping University, Linköping, Sweden

Email: karl@isy.liu.se

Abstract—Common data preprocessing routines often introduce considerable flaws in laser-based tracking of extended objects. As an alternative, extended target tracking methods, such as the Gamma-Gaussian-Inverse Wishart (GGIW) probability hypothesis density (PHD) filter, work directly on raw data. In this paper, the GGIW-PHD filter is applied to real world traffic scenarios. To cope with the large amount of data, a mixture clustering approach which reduces the combinatorial complexity and computation time is proposed. The effective segmentation of raw measurements with respect to spatial distribution and motion is demonstrated and evaluated on two different applications: pedestrian tracking from a vehicle and intersection surveillance.

I. INTRODUCTION

One of the major challenges in laser-based tracking for traffic perception is the deviation from the point target assumption. That is, the high angular resolution of the sensors usually yields a large amount of range measurements for individual objects and tracks can thus not be updated with a single measurement vector without preprocessing routines.

In such preprocessing steps, relevant objects are commonly detected using spatial segmentation and clustering [1]–[3] as well as shape fitting [4] or feature extraction [5] algorithms. However, these algorithms are heavily dependent on parametrization and suffer from over- or under-segmentation. Especially in changing environments, or in case of detecting different types of objects, it is nearly impossible to determine appropriate parameters. Since object tracking is directly influenced by the received data, errors during preprocessing heavily affect its reliability and accuracy.

To overcome the mentioned shortcomings, we suggest to use extended target tracking methods that work directly on raw data. The obtained tracks can either be used as output directly or serve as an improved input to classical tracking approaches. For instance, a separation between moving and stationary measurements can be used to select relevant readings for the subsequent tracking system. Also, the extended targets provide a partitioning of the measurements with additional velocity and orientation information which facilitates subsequent shape fitting or feature extraction.

The Gamma-Gaussian-Inverse Wishart (GGIW) probability hypothesis density (PHD) filter is a recently developed ex-

tended target PHD-filter which estimates an object's kinematic state, its extent and the number of measurements that it generates. The filter is an extension of the Gaussian-Inverse Wishart (GIW) -PHD filter [6], [7] which assumes a constant or state-dependent number of measurements. The additional measurement rate estimation in the GGIW-PHD-filter is done using the approach presented in [8].

In this work, the GGIW-PHD filter is applied for the first time to data from complex real-world scenarios with a large amount of range readings. In particular, the filter is tested on laser range finder data from two scenarios: a parking lot scene where the goal is to track moving pedestrians using multiple laser range finders on a stationary vehicle and an intersection scenario where several scanners were mounted above the intersection and provide a bird's-eye view. Here, the goal is to track all road users at the intersection.

Due to the large amount of measurements and the accompanying complexity, the update routine of the standard GGIW-PHD-filter requires considerable computation time. We thus propose a mixture clustering method to adapt the filter to real-world requirements, to improve filter performance and to significantly reduce computational costs.

The remainder of the paper is organized as follows: Section II presents the employed GGIW-PHD filter and Section III presents the proposed clustering approach to improve filtering performance. Experimental results are then discussed in Section IV and Section V concludes the paper.

II. THE GGIW-PHD FILTER

This section is intended to give an overview of the GGIW-PHD-filtering principle and presents modeling assumptions as well as prediction and update routines. Filter derivations are omitted due to page number constraints.

A. Models

1) *Object State*: The GGIW-PHD filter estimates an object's kinematic state, its extent and the number of measurements that it generates. The kinematic state in d -dimensional space is assumed to be distributed normally with mean $\mathbf{x}_k = [\mathbf{p}_k, \mathbf{v}_k]^T \in \mathbb{R}^{2d}$. It contains the position $\mathbf{p}_k \in \mathbb{R}^d$ and the velocity $\mathbf{v}_k \in \mathbb{R}^d$ of the object. Furthermore, an object

generates measurements according to a multivariate normal distribution centered at the object position with unknown covariance matrix $X_k \in \mathbb{S}_{++}^d$ [9], [10], where \mathbb{S}_{++}^n denotes the set of symmetric positive definite $n \times n$ -matrices. This design choice implies that the object has an elliptical shape. The amount of measurements that originate from the object is assumed to follow a Poisson distribution with an expected number of measurements $\gamma_k > 0$. The resulting state ξ_k of an extended object is thus given by

$$\xi_k \triangleq (\gamma_k, \mathbf{x}_k, X_k). \quad (1)$$

To estimate the components of the extended target state in a Bayesian framework, the joint probability density of the extended target state is given by the respective conjugate priors, in this case by a gamma-Gaussian-inverse Wishart distribution [8], [9]

$$p(\xi_k | \mathbf{Z}^k) = p(\gamma_k | \mathbf{Z}^k) p(\mathbf{x}_k | X_k, \mathbf{Z}^k) p(X_k | \mathbf{Z}^k) \quad (2a)$$

$$= \mathcal{G}(\gamma_k; \alpha_{k|k}, \beta_{k|k}) \mathcal{N}(\mathbf{x}_k; m_{k|k}, P_{k|k} \otimes X_k) \times \mathcal{IW}_d(X_k; v_{k|k}, V_{k|k}). \quad (2b)$$

which is conditioned on the previous sets of measurements \mathbf{Z}^k . Here, the Gamma probability distribution over the scalar $\gamma > 0$ is given by

$$\mathcal{G}(\gamma; \alpha, \beta) = \beta^\alpha \Gamma(\alpha)^{-1} \gamma^{\alpha-1} e^{-\beta\gamma}, \quad (3)$$

where $\Gamma(\cdot)$ is the gamma function, $\alpha > 0$ denotes the scalar shape parameter, and $\beta > 0$ is the scalar inverse scale parameter. Following [11, Definition 3.4.1], the inverse Wishart density function over the matrix $X \in \mathbb{S}_{++}^d$ is given by

$$\mathcal{IW}_d(X; v, V) = \frac{2^{-\frac{v-d-1}{2}} |V|^{\frac{v-d-1}{2}}}{\Gamma_d\left(\frac{v-d-1}{2}\right) |X|^{\frac{v}{2}}} \text{etr}\left(-\frac{1}{2} X^{-1} V\right), \quad (4)$$

where $v > 2d$ are the scalar degrees of freedom and $V \in \mathbb{S}_{++}^d$ is the parameter matrix. Moreover, $\text{etr}(\cdot) = \exp(\text{Tr}(\cdot))$ is the exponential of the matrix trace, \otimes denotes the Kronecker product, and $\Gamma_d(\cdot)$ denotes the multivariate gamma function. For notational convenience, (2) is abbreviated by

$$p(\xi_k | \mathbf{Z}^k) = \mathcal{GGIW}(\xi_k; \zeta_{k|k}), \quad (5)$$

in the following. The parameters of the GGIW density are given by $\zeta_{k|k} = \{\alpha_{k|k}, \beta_{k|k}, m_{k|k}, P_{k|k}, v_{k|k}, V_{k|k}\}$.

Note that the measurement rate is assumed to be independent of the object's kinematic and extension state in (2). However, the number of measurements per object definitely depends on the size and the distance of the object for most sensors in reality. This dependence is not modeled here, but the variance of the gamma distribution is sufficient to represent the variation of the number of measurements over time.

2) *Process Models*: The object's kinematic state is modeled using the process model [9]

$$\mathbf{x}_{k+1} = (F_{k+1|k} \otimes \mathbf{I}_d) \mathbf{x}_k + \mathbf{w}_{k+1}, \quad (6)$$

where \mathbf{I}_d is a $d \times d$ identity matrix. Furthermore, \mathbf{w}_{k+1} denotes the process noise which is a zero mean Gaussian distribution with covariance $\Delta_{k+1|k} = \mathbf{Q}_{k+1|k} \otimes X_{k+1}$ where [9]

$$\mathbf{Q}_{k+1|k} = \Sigma^2 \left(1 - e^{-2T_s/\theta}\right) \text{diag}([0 \ 1]), \quad (7)$$

Σ is the standard deviation of the velocity, T_s is the sampling time, and the maneuver correlation time is denoted by θ . The process matrix $F_{k+1|k}$ is given by [9]

$$F_{k+1|k} = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix}. \quad (8)$$

While the kinematic state changes over time, the measurement rate as well as the extension of the object are assumed to be approximately time-invariant.

3) *Measurement Model*: The measurement process is modeled using the linear measurement model [9]

$$\mathbf{z}_k = (H_k \otimes \mathbf{I}_d) \mathbf{x}_k + \mathbf{e}_k, \quad (9)$$

where $H_k = [1 \ 0]$ and the measurement noise \mathbf{e}_k follows a zero-mean Gaussian distribution whose covariance is given by the object's extension matrix X_k . Clutter measurements are included as Poisson process with intensity $\beta_{FA,k}$ and uniform spatial distribution over the surveillance area volume $V(\mathcal{A})$.

B. Prediction

The posterior PHD $D_{k|k}(\cdot)$ at time k is given by a mixture of GGIW distributions

$$D_{k|k}(\xi_k) = \sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)} \mathcal{GGIW}(\xi_k; \zeta_{k|k}^{(j)}) \quad (10)$$

with $w_{k|k}^{(j)}$ and $\zeta_{k|k}^{(j)}$ denoting the weight and the GGIW density parameters of component j , respectively. $J_{k|k}$ represents the number of GGIW components and the expected amount of extended objects is given by $\sum_{j=1}^{J_{k|k}} w_{k|k}^{(j)}$.

The prior at time step $k+1$ is obtained by prediction of (10) which again results in a GGIW mixture

$$D_{k+1|k}(\xi_{k+1}) = D_{k+1|k}^b(\xi_{k+1}) + D_{k+1|k}^s(\xi_{k+1}). \quad (11)$$

It is composed of the intensity of new-born objects $D_{k+1|k}^b(\xi_{k+1})$ and the intensity of persisting objects $D_{k+1|k}^s(\xi_{k+1})$. New-born objects are modeled by the following PHD intensity, using a uniform birth intensity for the position [12],

$$D_k^b(\xi_k) = w_k^{(b)} \mathcal{U}(\mathbf{p}_k) \mathcal{N}(\mathbf{v}_k; m_k^{(b)}, P_k^{(b)} \otimes X_k) \times \mathcal{G}(\gamma_k; \alpha_k^{(b)}, \beta_k^{(b)}) \mathcal{IW}_d(X_k; v_k^{(b)}, V_k^{(b)}). \quad (12)$$

The respective birth parameters have to be chosen in dependence on the application.

The PHD intensity of the persisting objects is given by

$$D_{k+1|k}^s(\xi_{k+1}) = \sum_{j=1}^{J_{k|k}} w_{k+1|k}^{(j)} \mathcal{GGIW}(\xi_{k+1}; \zeta_{k+1|k}^{(j)}), \quad (13)$$

where the prior component weights $w_{k+1|k}^{(j)} = P_S w_{k|k}^{(j)}$ are reduced by the persistence probability P_S . The mean and the covariance of the kinematic state are predicted using

$$m_{k+1|k}^{(j)} = (F_{k+1|k} \otimes \mathbf{I}_d) m_{k|k}^{(j)}, \quad (14a)$$

$$P_{k+1|k}^{(j)} = F_{k+1|k} P_{k|k}^{(j)} F_{k+1|k}^T + \mathbf{Q}_{k+1|k}. \quad (14b)$$

As mentioned above, the number of measurements per object as well as the object extent are expected to be approximately constant. Hence, the parameters of the Gamma and inverse Wishart distributions are predicted by

$$\alpha_{k+1|k}^{(j)} = \frac{\alpha_{k|k}^{(j)}}{\eta_k}, \quad \beta_{k+1|k}^{(j)} = \frac{\beta_{k|k}^{(j)}}{\eta_k}, \quad (15a)$$

$$v_{k+1|k}^{(j)} = 2d + 2 + e^{-T_s/\tau} (v_{k|k}^{(j)} - 2d - 2), \quad (15b)$$

$$V_{k+1|k}^{(j)} = \frac{v_{k+1|k}^{(j)} - 2d - 2}{v_{k|k}^{(j)} - 2d - 2} V_{k|k}^{(j)}. \quad (15c)$$

This corresponds to keeping the expected value constant and increasing the variance.

C. Measurement Update

At time step k , the prior PHD intensity (11) can be rewritten as

$$D_{k|k-1}(\xi_k) = \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} \mathcal{GGIW}(\xi_k; \zeta_{k|k-1}^{(j)}). \quad (16)$$

Then, the posterior PHD is again a GGIW mixture

$$D_{k|k}(\xi_k) = D_{k|k}^m(\xi_k) + D_{k|k}^b(\xi_k) + D_{k|k}^d(\xi_k) \quad (17)$$

which comprises the intensities for missed detections, new-born objects, and detected objects.

The updated intensities for new-born objects and for the detection of persisting objects are obtained using the measurement set $\mathbf{Z}_k = \{\mathbf{z}_k^{(j)}\}_{j=1}^{N_{z,k}}$. All sensor readings are separated into distinct cells \mathbf{W} under the constraint that the union of all cells equals the measurement set \mathbf{Z} . Intuitively, each cell contains measurements that are assumed to originate from a single object. The resulting cells then constitute a possible partition \mathcal{P} of all readings. For a cell \mathbf{W} , the centroid measurement and scatter matrix are given by

$$\bar{\mathbf{z}}_k^{\mathbf{W}} = \frac{1}{|\mathbf{W}|} \sum_{\mathbf{z}_k^{(i)} \in \mathbf{W}} \mathbf{z}_k^{(i)}, \quad (18a)$$

$$\mathbf{Z}_k^{\mathbf{W}} = \sum_{\mathbf{z}_k^{(i)} \in \mathbf{W}} (\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{\mathbf{W}}) (\mathbf{z}_k^{(i)} - \bar{\mathbf{z}}_k^{\mathbf{W}})^T. \quad (18b)$$

In order to ensure Bayes optimality, all possible partitions of \mathbf{Z}_k should be used to update the intensities. However, using

all partitions is computationally intractable due to combinatorial complexity [6], [13], [14]. Section III will thus discuss methods to reduce the computational burden.

Since the measurement update significantly increases the number of components $J_{k|k}$, the intensity has to be truncated using mixture pruning and merging [8], [15] after the update step.

1) *Missed detections*: The detection probability of a persisting object is given by $P_{k,D}^e = (1 - e^{-\gamma_k}) P_D$ where $P_D \in (0, 1)$ [16]. As in [12], the detection probability of new born objects is assumed to be $P_{k,D}^e = 1$. Hence, the missed detection intensity only contains components originating from persisting objects. The updated PHD for not detected previously existing objects at time step k is approximated as

$$\begin{aligned} D_{k|k}^m(\xi_k) &= \sum_{j=1}^{J_{k|k-1}} (1 - P_{k,D}^e) w_{k|k-1}^{(j)} \mathcal{GGIW}(\xi_k; \zeta_{k|k-1}^{(j)}) \\ &\approx \sum_{j=1}^{J_{k|k-1}} \tilde{w}_{k|k}^{(j)} \mathcal{GGIW}(\xi_k; \tilde{\zeta}_{k|k}^{(j)}), \end{aligned} \quad (19a)$$

where the parameters of the GGIW density are given by

$$\tilde{\zeta}_{k|k}^{(j)} = \left\{ \tilde{\alpha}_{k|k}^{(j)}, \tilde{\beta}_{k|k}^{(j)}, m_{k|k-1}^{(j)}, P_{k|k-1}^{(j)}, v_{k|k-1}^{(j)}, V_{k|k-1}^{(j)} \right\}. \quad (20)$$

Computing the parameters $\tilde{\alpha}_{k|k}^{(j)}$ and $\tilde{\beta}_{k|k}^{(j)}$ of the Gamma distribution, as well as the weights $\tilde{w}_{k|k}^{(j)}$, requires using the gamma-mixture reduction proposed in [8].

2) *Detection of new born objects*: The measurement update of the birth intensity (12) is given by

$$D_{k|k}^b(\xi_k) = \sum_{\mathcal{P} \angle \mathbf{Z}_k} \sum_{\mathbf{W} \in \mathcal{P}} w_{k|k}^{(b,\mathbf{W})} \mathcal{GGIW}(\xi_k; \zeta_{k|k}^{(b,\mathbf{W})}), \quad (21)$$

where $\mathcal{P} \angle \mathbf{Z}$ represents all possible partitions \mathcal{P} of the set \mathbf{Z} . The kinematic state is updated using

$$m_{k|k}^{(b,\mathbf{W})} = \left[(\bar{\mathbf{z}}_k^{\mathbf{W}})^T, (m_k^{(b)})^T \right]^T, \quad (22a)$$

$$P_{k|k}^{(b,\mathbf{W})} = \text{blkdiag}(|\mathbf{W}|^{-1}, P_k^{(b)}), \quad (22b)$$

the updated parameters for the measurement rate follow

$$\alpha_{k|k}^{(b,\mathbf{W})} = \alpha_k^{(b)} + |\mathbf{W}|, \quad \beta_{k|k}^{(b,\mathbf{W})} = \beta_k^{(b)} + 1, \quad (23)$$

and the Wishart parameters for the object's extension are given

$$v_{k|k}^{(b,\mathbf{W})} = v_{k|k-1}^{(b)} + |\mathbf{W}|, \quad (24a)$$

$$V_{k|k}^{(b,\mathbf{W})} = V_{k|k-1}^{(b)} + \mathbf{Z}_k^{\mathbf{W}}. \quad (24b)$$

Equations for the update of the weight $w_{k|k}^{(b,\mathbf{W})}$ will be given in Section II-C4.

3) *Detection of persisting objects*: The update of the PHD of persisting objects (13) using the measurement set \mathbf{Z}_k is given by

$$D_{k|k}^d(\xi_k) = \sum_{\mathcal{P} \angle \mathbf{Z}_k} \sum_{\mathbf{W} \in \mathcal{P}} \sum_{j=1}^{J_{k|k-1}} w_{k|k}^{(j,\mathbf{W})} \mathcal{GGIW}(\xi_k; \zeta_{k|k}^{(j,\mathbf{W})}). \quad (25)$$

The kinematic state of the objects is updated using [9]

$$m_{k|k}^{(j, \mathbf{W})} = m_{k|k-1}^{(j)} + \left(K_{k|k-1}^{(j, \mathbf{W})} \otimes \mathbf{I}_d \right) \varepsilon_{k|k-1}^{(j, \mathbf{W})}, \quad (26a)$$

$$P_{k|k}^{(j, \mathbf{W})} = P_{k|k-1}^{(j)} - K_{k|k-1}^{(j, \mathbf{W})} S_{k|k-1}^{(j, \mathbf{W})} \left(K_{k|k-1}^{(j, \mathbf{W})} \right)^T, \quad (26b)$$

$$S_{k|k-1}^{(j, \mathbf{W})} = H_k P_{k|k-1}^{(j)} H_k^T + \frac{1}{|\mathbf{W}|}, \quad (26c)$$

$$K_{k|k-1}^{(j, \mathbf{W})} = P_{k|k-1}^{(j)} H_k^T \left(S_{k|k-1}^{(j, \mathbf{W})} \right)^{-1}, \quad (26d)$$

where $S_{k|k-1}^{(j, \mathbf{W})}$ is the scalar innovation factor and $K_{k|k-1}^{(j, \mathbf{W})}$ is the gain vector. Identical to the detection of new born objects, the parameters of the gamma distribution are updated using

$$\alpha_{k|k}^{(j, \mathbf{W})} = \alpha_{k|k-1}^{(j)} + |\mathbf{W}|, \quad \beta_{k|k}^{(j, \mathbf{W})} = \beta_{k|k-1}^{(j)} + 1. \quad (27)$$

Finally, the measurement updated object extension is given by

$$v_{k|k}^{(j, \mathbf{W})} = v_{k|k-1}^{(j)} + |\mathbf{W}|, \quad (28a)$$

$$V_{k|k}^{(j, \mathbf{W})} = V_{k|k-1}^{(j)} + N_{k|k-1}^{(j, \mathbf{W})} + Z_k^{\mathbf{W}}, \quad (28b)$$

$$\varepsilon_{k|k-1}^{(j, \mathbf{W})} = \bar{\mathbf{z}}_k^{\mathbf{W}} - (H_k \otimes \mathbf{I}_d) m_{k|k-1}^{(j)}, \quad (28c)$$

$$N_{k|k-1}^{(j, \mathbf{W})} = \left(S_{k|k-1}^{(j, \mathbf{W})} \right)^{-1} \varepsilon_{k|k-1}^{(j, \mathbf{W})} \left(\varepsilon_{k|k-1}^{(j, \mathbf{W})} \right)^T, \quad (28d)$$

where $\varepsilon_{k|k-1}^{(j, \mathbf{W})}$ is the innovation vector and $N_{k|k-1}^{(j, \mathbf{W})}$ denotes the innovation covariance.

4) *Component weights*: The weights of the measurement updated GGIW components for new born and persisting objects are

$$w_{k|k}^{(b, \mathbf{W})} = \frac{\omega_{\mathcal{P}}}{d_{\mathbf{W}} \beta_{FA,k}^{|\mathbf{W}|}} \mathcal{L}_k^{(b, \mathbf{W})} w_k^{(b)} \frac{1}{V(\mathcal{A})} \quad (29a)$$

$$w_{k|k}^{(j, \mathbf{W})} = \frac{\omega_{\mathcal{P}} P_D}{d_{\mathbf{W}} \beta_{FA,k}^{|\mathbf{W}|}} \mathcal{L}_k^{(j, \gamma)} \mathcal{L}_k^{(j, \mathbf{x}, X)} w_{k|k-1}^{(j)}, \quad (29b)$$

with

$$\omega_{\mathcal{P}} = \frac{\prod_{\mathbf{W} \in \mathcal{P}} d_{\mathbf{W}}}{\sum_{\mathcal{P}' \subset \mathcal{Z}_k} \prod_{\mathbf{W}' \in \mathcal{P}'} d_{\mathbf{W}'}} \quad (29c)$$

$$d_{\mathbf{W}} = \delta_{|\mathbf{W}|,1} + \frac{\mathcal{L}_k^{(b, \mathbf{W})} w_k^{(b)}}{\beta_{FA,k}^{|\mathbf{W}|} V(\mathcal{A})} + \sum_{j=1}^{J_{k|k-1}} \frac{P_D \mathcal{L}_k^{(j, \mathbf{W})} w_{k|k-1}^{(j)}}{\beta_{FA,k}^{|\mathbf{W}|}} \quad (29d)$$

$$\begin{aligned} \mathcal{L}_k^{(b, \mathbf{W})} &= \frac{|\mathbf{W}|^{-d/2}}{\pi^{|\mathbf{W}|(d-1)/2}} \frac{\Gamma_d \left(\frac{v_{k|k}^{(b, \mathbf{W})}}{2} \right) \left| V_k^{(b, \mathbf{W})} \right|^{\frac{v_{k|k}^{(b)} - d - 1}{2}}}{\Gamma_d \left(\frac{v_k^{(b)}}{2} \right) \left| V_k^{(b, \mathbf{W})} \right|^{\frac{v_{k|k}^{(b, \mathbf{W})} - d - 1}{2}}} \\ &\quad \times \frac{\Gamma \left(\alpha_{k|k}^{(b, |\mathbf{W}|)} \right) \left(\beta_{k|k}^{(b)} \right)^{\alpha_k^{(b)}}}{\Gamma \left(\alpha_k^{(b)} \right) \left(\beta_{k|k}^{(b, \mathbf{W})} \right)^{\alpha_{k|k}^{(b, |\mathbf{W}|)}}} \end{aligned} \quad (29e)$$

$$\mathcal{L}_k^{(j, \mathbf{W})} = \frac{(\pi^{|\mathbf{W}|} |\mathbf{W}|)^{-\frac{d}{2}} \left| V_{k|k-1}^{(j)} \right|^{\frac{v_{k|k-1}^{(j)}}{2}} \Gamma_d \left(\frac{v_{k|k}^{(j, \mathbf{W})}}{2} \right)}{\left(S_{k|k-1}^{(j, \mathbf{W})} \right)^{\frac{d}{2}} \left| V_{k|k}^{(j, \mathbf{W})} \right|^{\frac{v_{k|k}^{(j, \mathbf{W})}}{2}} \Gamma_d \left(\frac{v_{k|k-1}^{(j)}}{2} \right)}$$

$$\times \frac{\Gamma \left(\alpha_{k|k}^{(j, |\mathbf{W}|)} \right) \left(\beta_{k|k-1}^{(j)} \right)^{\alpha_{k|k-1}^{(j)}}}{\Gamma \left(\alpha_{k|k-1}^{(j)} \right) \left(\beta_{k|k}^{(j, \mathbf{W})} \right)^{\alpha_{k|k}^{(j, |\mathbf{W}|)}}} \quad (29f)$$

III. PARTITIONING AND CLUSTERING THE GGIW PHD

As mentioned in Section II, the update of the PHD with all possible partitions is computationally intractable. Different methods to find only the likely partitions have been proposed in [6], [14]. Partitions are obtained by using distance-based methods which separate measurements according to varying distance thresholds, grouping points based on the position of predicted targets, and expectation maximization methods.

Especially distance based clustering methods suffer from the drawback that a single threshold is often not the best choice for all objects in the scene. Figure 1 and Figure 2 illustrate the influence of the partitioning threshold for data recorded at the intersection. While a threshold of $d = 0.25$ m obtains the correct partitioning for the pedestrian and the bicycle, the same threshold returns a huge amount of cells for the car. In contrast, $d = 0.5$ m delivers only a single cell for the pedestrian and the bicycle but partitions the car measurements correctly into a single cell.

To overcome such issues, a sub-partitioning routine which detects possible under-segmentation and reruns the distance-based segmentation for the corresponding cell was suggested in [6]. Additional partitions are then obtained by replacing the affected cells with the smaller propositions. As noted in [6], however, this method does not consider the occurrence of multiple cases of under segmentation since this would again lead to combinatorial complexity.

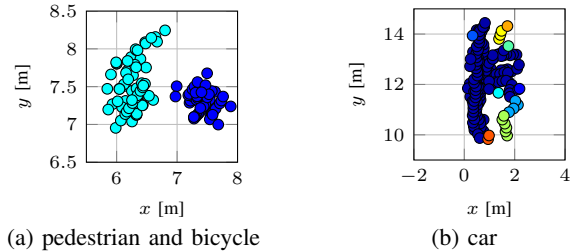


Fig. 1: Partitioning results for distance $d = 0.25$ m. The obtained results are illustrated using color coding.

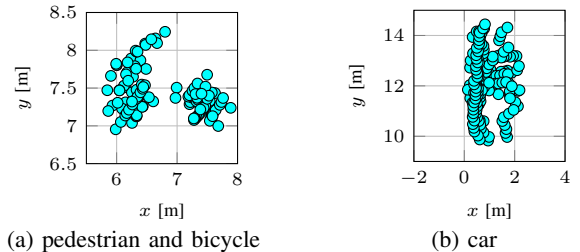


Fig. 2: Partitioning results for distance $d = 0.5$ m. The obtained results are illustrated using color coding.

In this paper we take an alternative approach towards achieving improved tracking results at a reasonable computational complexity. Given a partition \mathcal{P} of a new set of

measurements \mathbf{Z}_k and a predicted PHD intensity $D_{k|k-1}(\cdot)$, many of the cell and GGIW component pairs will be distant, and the effect of updating that part of the PHD intensity with that cell will be negligible. We propose to use this to alleviate some of the computational complexity of the GGIW PHD filter.

The approach has the following steps:

- 1) Cluster the predicted persisting object GGIW components, cf. (13), into N_c mutually exclusive mixtures,

$$\begin{aligned} D_{k|k-1}^s(\xi_k) &= \sum_{j=1}^{J_{k|k-1}} w_{k|k-1}^{(j)} \mathcal{GGIW}(\xi_k; \zeta_{k|k-1}^{(j)}) \\ &= \sum_{i=1}^{N_c} \sum_{j \in \mathbb{J}^{(i)}} w_{k|k-1}^{(j)} \mathcal{GGIW}(\xi_k; \zeta_{k|k-1}^{(j)}), \\ &= \sum_{i=1}^{N_c} D_{k|k-1}^{s, \mathbb{J}^{(i)}}(\xi_k), \end{aligned} \quad (30)$$

where $\mathbb{J}^{(i)}$ enumerates the component indices in the i th cluster, $\mathbb{J}^{(n)} \cap \mathbb{J}^{(m)} = \emptyset$ if $n \neq m$, and

$$\bigcup_{i=1}^{N_c} \mathbb{J}^{(i)} = \mathbb{J} = \{1, \dots, J_{k|k-1}\} \quad (31)$$

- 2) From the full set of measurements \mathbf{Z}_k , assign only the relevant measurements to each cluster, denoted $\mathbf{Z}_k^{\mathbb{J}^{(i)}}$. For each cluster, compute $D_{k|k}^{d, \mathbb{J}^{(i)}}(\xi_k)$, cf. (25), by updating the intensity $D_{k|k-1}^{s, \mathbb{J}^{(i)}}(\xi_k)$ using the measurements $\mathbf{Z}_k^{\mathbb{J}^{(i)}}$ rather than the full measurement set \mathbf{Z}_k .
- 3) The PHD intensity corresponding to detected persisting objects is approximated as

$$D_{k|k}^d(\xi_k) \approx \sum_{i=1}^{N_c} D_{k|k}^{d, \mathbb{J}^{(i)}}(\xi_k), \quad (32)$$

In the second step above, multiple partitions $\mathcal{P}^{\mathbb{J}^{(i)}}$ of each set $\mathbf{Z}_k^{\mathbb{J}^{(i)}}$ are considered. In the measurement update step, one of the partitions will typically dominate remaining partitions, i.e. have a significantly higher weight $\omega_{\mathcal{P}}$, cf. (29c). Assigning only the relevant measurements to the mixture clusters allows the measurement partitions $\mathcal{P}^{\mathbb{J}^{(i)}}$ to be determined independently. This way the clusters $\mathbb{J}^{(i)}$ can have dominating partitions $\mathcal{P}^{\mathbb{J}^{(i)}}$ that correspond to different distance thresholds d . This effectively solves the problem that is highlighted in Figures 1 and 2.

Note that the same effect can be achieved by considering more partitions \mathcal{P} for the full PHD intensity, i.e. without the clustering presented above. However, the number of measurement cells and mixture components is comparatively smaller for each cluster compared to the full PHD. Thus, using the proposed clustering procedure the amount of update combinations is reduced significantly, and the procedure achieves a considerable computation speed-up without a significant performance sacrifice.

The GGIW component clustering is performed as follows:

- 1) A partition \mathcal{P}_{max} of the measurement set \mathbf{Z}_k is obtained using a maximum distance d_{max} . An appropriate choice for d_{max} ensures that the measurements of two cells \mathbf{W}_m and \mathbf{W}_n of the partition \mathcal{P}_{max} are highly unlikely to belong to the same extended object.
- 2) The spatial proximity of the cells $\mathbf{W}_i \in \mathcal{P}$ and the GGIW components $\mathcal{GGIW}(\xi_{k+1}; \zeta_{k+1|k}^{(j)})$ are computed using

$$d^{(j)}(\mathbf{W}_i) = \min_{\mathbf{z} \in \mathbf{W}_i} \left(\left(d^{(j)}(\mathbf{z}) \right)^T R_k^{-1} \left(d^{(j)}(\mathbf{z}) \right) \right) \quad (33a)$$

$$d^{(j)}(\mathbf{z}) = \mathbf{z} - \tilde{H}_k m_{k+1|k}^{(j)} \quad (33b)$$

$$R_k = \hat{X}_{k|k-1}^{(j)} + \tilde{H}_k \hat{P}_{k|k-1}^{(j)} \tilde{H}_k^T \quad (33c)$$

$$\hat{X}_{k|k-1}^{(j)} = V_{k|k-1}^{(j)} \left(v_{k|k-1}^{(j)} - 2d - 2 \right)^{-1} \quad (33d)$$

$$\hat{P}_{k|k-1}^{(j)} = P_{k|k-1}^{(j)} \otimes \hat{X}_{k|k-1}^{(j)} \quad (33e)$$

$$\tilde{H}_k = H_k \otimes \mathbf{I}_d \quad (33f)$$

- 3) Using a standard gating procedure, see e.g. [17], the distance $d^{(j)}(\mathbf{W}_i)$ is used to determine the contributing measurement cells for each GGIW component j .
- 4) Let $\mathbb{W} = \{1, \dots, |\mathcal{P}_{max}|\}$ enumerate the cells in the partition \mathcal{P}_{max} . A GGIW cluster $\mathbb{J}^{(i)}$ is given by the GGIW components that share at least one measurement cell \mathbf{W}_i . This way N_c groupings

$$\mathcal{G}^{(n)} = (\mathbb{J}^{(n)}, \mathbb{W}^{(n)}) \quad (34)$$

are obtained, for which it holds $\mathbb{J}^{(n)} \cap \mathbb{J}^{(m)} = \emptyset$ and $\mathbb{W}^{(n)} \cap \mathbb{W}^{(m)} = \emptyset$ if $n \neq m$, and where

$$\mathbf{Z}_k^{\mathbb{J}^{(i)}} = \bigcup_{j \in \mathbb{W}^{(i)}} \mathbf{W}_j. \quad (35)$$

In case of a sufficiently large gating threshold (in our implementation $d^{(j)}(\mathbf{W}_i) > 9$), the GGIW partitions $D_{k|k-1}^{s, \mathbb{J}^{(n)}}$ have negligible influence on each other in the measurement update, because the measurement likelihood for measurements of other cells $\mathbb{W}^{(m)}$, $n \neq m$, is close to zero. The proof of this property is straight-forward using the derivation of the measurement likelihood (29f) [6, Appendix A].

IV. EXPERIMENTS

To apply the GGIW-PHD-filter to the pedestrian tracking and intersection scenarios, it was implemented in MATLAB. A distance-based segmentation method [14] which automatically determines possible distance thresholds is applied to the measurements of each cluster. It provides several feasible partitions. Moreover, a simple labelling method which assigns unique identifiers to mixture components and their successors was added to be able to obtain continuous track trajectories. For generating a list of tracks from the GGIW mixture distribution, each mixture component with a weight $w_{k|k}^{(j)} > 0.5$ is selected and stored as distinct track.

In the following, the two experimental scenarios, pedestrian tracking and the urban intersection, as well as the respective tracking results are presented.



Fig. 3: Pedestrian tracking scenario with two persons walking on a parking lot

A. Pedestrian Tracking Scenario

The first scenario is used to demonstrate how the filter behaves in the presence of clutter and how well it is able to separate relevant moving objects from stationary tracks. In particular, we want to distinguish between moving pedestrians and parked vehicles.

1) *Experimental Set-Up*: Two persons were recorded from a stationary vehicle while walking on a parking lot as depicted in Fig. 3. The first pedestrian enters the field of view next to the experimental vehicle and finally leaves it at a distance of 25m. As soon as the first pedestrian made half of its way, a second person enters, walks towards the vehicle and finally does a small loop.

The experimental vehicle is equipped with three Ibeo Lux laser range finders which provide measurements in four layers and which are mounted in the center and corners of the front bumper. Ground readings are automatically detected and removed by the sensors. In the present scenario, this results in an average of 1160 range measurements per time step.

2) *Results*: Figure 4 illustrates the obtained tracking results. Tracks are represented by their two-sigma ellipses. For clarity, the trajectories are only plotted for the two pedestrians which are continuously tracked by the filter. Exemplary ellipses with corresponding measurements are plotted for every 12th time step to demonstrate how the filter adapts the object size to the changing appearance of the pedestrians. This is especially evident for the blue track as it gets close to the sensors. The ellipses follow the pendulum motion of the legs. That is, they are rather compressed in direction of motion as the two legs are close-by and are considerably elongated as the pedestrian makes a big step.

The green ellipses represent tracks originating from parked vehicles and vegetation. Since they are correctly identified as static objects in most time steps and thus do not exhibit significant motion, trajectories are omitted and all ellipses are taken from a single and exemplary time step. Due to the chosen partitioning settings, i.e. distance thresholds from 0.5 m to 1 m, the tracking system tends to explain close-by measurements by relatively large ellipses which sometimes leads to merged objects as for the track at approximately (-6m, 10m). Also, note how the elliptic profile has difficulties to represent the distinct L-shape at (3m, 4m) and how the ellipse is focused at the measurement center, thus treating sparse and distant readings as outliers. Still, all the information needed to classify the object as static is provided.

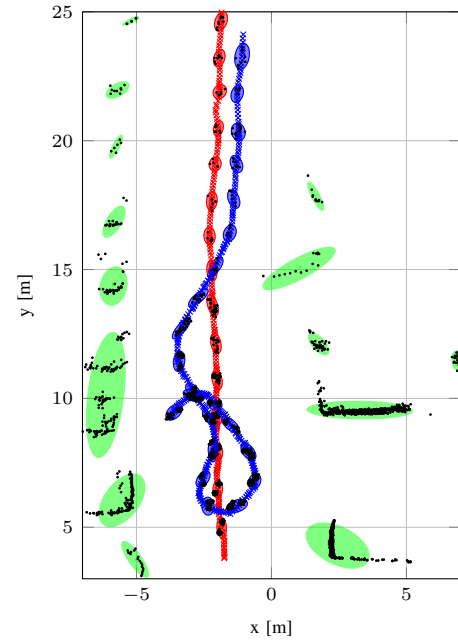


Fig. 4: Pedestrian tracking results: the green ellipses are stationary tracks taken from a single time step, the red and blue trajectories illustrate the pedestrian tracks with exemplary ellipses and corresponding measurements for every 12th time step

To evaluate the classification performance for stationary and moving objects, the number of identified moving objects is compared to the true value in Fig. 5. Here, the amount of moving objects is simply obtained by counting the mixture components with an estimated speed greater than 0.5 m/s. Apart from a few deviations, the number is estimated correctly in 93.3% of the time steps. The remaining aberrations can be explained by three causes. Firstly, the entering pedestrians are not detected as moving objects instantly. This is aggravated by the fact that the measurement centroids do not exhibit considerable motion as the objects enter the field of view and expand to their actual size. Secondly, the weights of the mixture component representing the first pedestrian drop below a value of 0.5 at 12.24s, 12.72s, and 15.76s. Since this threshold is used for determining relevant mixture components, they are not counted as moving objects even though the track still exists. The decreasing weight is caused by missing measurements due to occlusion or a large distance to the sensor, respectively. The third effect causes an overestimated number of objects as illustrated by the pronounced spikes. Here, some of the static object tracks are subject to phantom motion which is generated by a change of object appearance. These changes are a consequence of objects being partially occluded by the pedestrians. However, the rather large extent of the static objects makes the tracks rather insensitive to small changes and hence, phantom motion occurs rarely.

3) *Performance*: The average computation time for each time step using the MATLAB implementation is 0.76 seconds on an i7-3520M processor with 2.9 GHz. Due to the relatively

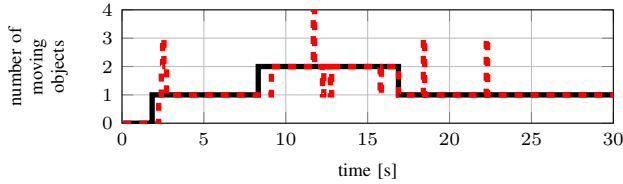


Fig. 5: Number of identified moving objects (dashed red) in comparison to the true value (black). In 93.3% of the time steps, the estimate is correct.

small interval of possible distance thresholds, partitioning the measurement clusters is relatively clear in many cases and a single partitioning proposition often suffices. When taking into account the possible combinations, however, the filter considers 18.4 variants in average. In contrast, the implementation of the standard GGIW-PHD filter without clustering requires a computation time of 1.88 seconds per time step while at the same time only considering an average of 5.1 partitions.

Since there is no deliberate parallelization, the computation is mostly performed on a single core. By optimizing the code, exploiting the potential for parallel processes due to the clustering approach and by porting the system to low level programming languages, a considerable speed-up is expectable.

B. Intersection Scenario

From the first scenario it was observable that elliptical track shapes are a rather precise approximation for pedestrian appearance, but have difficulties to represent rectangular objects such as vehicles. In this subsection, the GGIW-PHD-filter is applied to the intersection scenario. Apart from tracking road users of different kind, i.e. a vehicle, a pedestrian and a bicycle, we now demonstrate how the filter is able to improve subsequent feature extraction routines.

1) *Experimental Set-Up:* To observe road users, a laser range finder network consisting of multiple 8-layer SICKLDMRS scanners has been installed at a public intersection in Aschaffenburg, Germany, within the joint project Ko-PER of the research initiative Ko-FAS [18]. The sensors are mounted at least 5 m above the ground at lamp posts or traffic lights and provide a bird's-eye view of the current intersection scene. In Fig. 6 a realistic 3D model of the intersection together with the point cloud gathered by the laser range finders is depicted. For further details on the intersection set-up refer to [19]. Since the mounting positions of the sensors are static, it is feasible to train a background model to remove all measurements originating from static objects like buildings or street surface. The remaining point cloud, which only represents the road users at the intersection, is provided to the GGIW-PHD filter.

Since the elliptical shape is not a suitable representation for rectangular vehicles, the points within the two-sigma ellipse of an estimated track are used to determine a bounding box which is a common representation of vehicles and can be easily compared to available ground truth data. The bounding box is centered at the estimated position and oriented according to the estimated direction of motion. Subsequently, the length

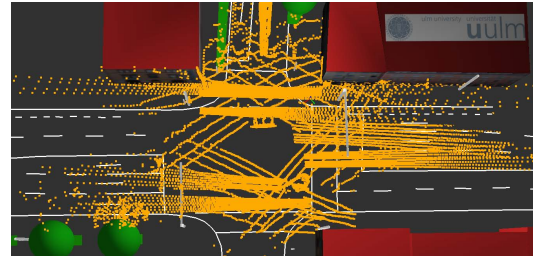


Fig. 6: Realistic 3D model of the Ko-PER intersection. The aligned range measurements of the 14 laserscanners are depicted by orange points.

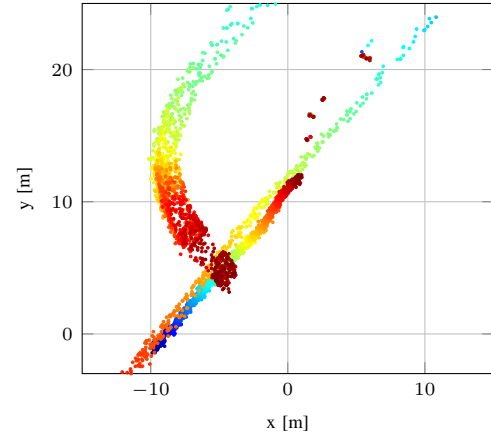


Fig. 7: Evolution of background subtracted laserscanner measurements in time. Time is color coded from blue to red. For purpose of illustration the plot is rotated through ca. 125 degrees compared to Fig. 6.

and width of the box are determined based on the measurement points. By using orientation information from raw data tracking, valuable information is introduced to this feature extraction routine which facilitates the process considerably.

Tracking performance evaluation of the filter is based on a sequence in which one pedestrian, one cyclist, and one car maneuver through the intersection. Figure 7 illustrates the content of the sequence. The pedestrian appears at the lower left corner of Fig. 7 and crosses the street illustrated in the upper part of Fig. 6 while the cyclist crosses this street in the opposite direction. The car performs a left turn maneuver and approaches with a velocity of approx. 12 m/s.

2) *Results:* The tracking results depicted by Fig. 8 show continuous tracks for all of the three objects on their way through the intersection. Even during the pass-by of the pedestrian and the cyclist, the tracks are not lost and no ID switches occur. To create a meaningful illustration of the tracking results, estimated bounding boxes are plotted for the car at every 15th and for the pedestrian and cyclist at every 5th time step. Since highly accurate reference data is available for the car trajectory, Fig. 8 additionally shows the true oriented bounding boxes for the car. While the vehicle is approaching the intersection, the laser range finder network is not able to detect the entire vehicle extent due to the sensor field of view. Therefore, the estimation of the dimension can just be

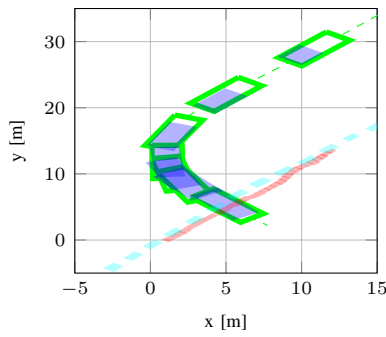


Fig. 8: Tracking results with estimated oriented bounding boxes of the road users. Each track is visualized with a different color to show continuity. Reference trajectory and bounding box of the car (blue) are depicted in green.

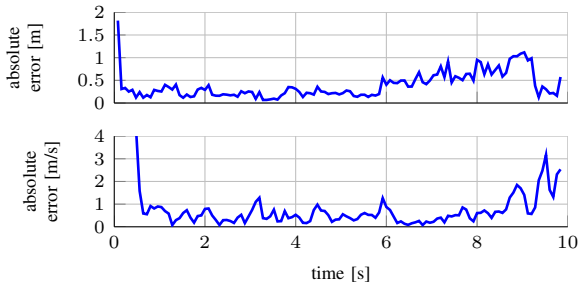


Fig. 9: Absolute error between the car's estimated position and velocity and their reference.

evaluated in the central region of the intersection where the estimated and the true bounding box match precisely.

As observable from the matching front edges of the reference and estimated bounding boxes, the position of the car is estimated accurately during the whole driving maneuver. This impression is confirmed by the absolute error of the position and velocity plotted in Fig. 9. To calculate the absolute error, the estimated state as well as its reference have to be transformed to a common reference point in every time step. For the left turning maneuver, the reference point is chosen to be the center of the car's front, since this point can be detected continuously by the laser range finder network. Except for the period from 6 s to 9.5 s, the error in the position is below 0.5 m. In the mentioned period, the car moves very slow. Thus, it is challenging to accurately estimate its velocity. This directly affects the determined bounding box, since the orientation of the box is based on the estimated motion direction. Furthermore, the error of the velocity does not exceed 1.2 m/s except at the end of the sequence, where the car leaves the perception area of the system. Keeping in mind that the absolute error of the position and the velocity incorporates the errors in orientation and extent estimation due to the reference point transformation, the GGIW-PHD-filter achieves a very accurate road user tracking which does not require any error prone and heuristic preprocessing.

V. CONCLUSION

In the present work, we applied the GGIW-PHD-filter to two different data sets with a considerable amount of mea-

surements. By introducing a mixture clustering procedure, we were able to separate the filtering procedure into independent problems thereby reducing the combinatorial complexity and computation time considerably. The experimental results show that the filter works on real world data, is able to track extended objects precisely and can also be used to improve a subsequent feature extraction routine. The proposed clustering method allows for a real-time capable implementation using parallelization. Using the filter in a non-stationary sensor set-up, i.e. from a moving vehicle is part of our future intentions.

REFERENCES

- [1] M. Himmelsbach, F. v. Hundelshausen, and H. Wuensche, "Fast segmentation of 3d point clouds for ground vehicles," in *IEEE Intelligent Vehicles Symposium (IV)*, June 2010, pp. 560–565.
- [2] D. Meissner, S. Reuter, and K. Dietmayer, "Combining the 2d and 3d world: A new approach for point cloud based object detection," in *Intelligent Signal Processing Conference*, 12 2013.
- [3] C. Premebida and U. Nunes, "Segmentation and geometric primitives extraction from 2d laser range data for mobile robot applications," in *Actas do Encontro Científico Robótica 2005*, 2005, pp. 17–25.
- [4] M. Munz, "Generisches sensorfusionsframework zur gleichzeitigen zustands- und existenzschätzung für die fahrzeugumfeldererkennung," Ph.D. dissertation, Ulm University, 7 2011.
- [5] V. Nguyen, A. Martinelli, N. Tomatis, and R. Siegwart, "A comparison of line extraction algorithms using 2d laser rangefinder for indoor mobile robotics," in *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2005, pp. 1929–1934.
- [6] K. Granström and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 5657–5671, Nov. 2012.
- [7] —, "Implementation of the GIW-PHD filter," Department of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden, Tech. Rep. LiTH-ISY-R-3046, Mar. 2012. [Online]. Available: <http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-94585>
- [8] —, "Estimation and Maintenance of Measurement Rates for Multiple Extended Target Tracking," in *Proceedings of the International Conference on Information Fusion*, Singapore, Jul. 2012, pp. 2170–2176.
- [9] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 1042–1059, Jul. 2008.
- [10] M. Feldmann, D. Fränken, and J. W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, Apr. 2011.
- [11] A. K. Gupta and D. K. Nagar, *Matrix variate distributions*, ser. Chapman & Hall/CRC monographs and surveys in pure and applied mathematics. Chapman & Hall, 2000.
- [12] M. Beard, B.-T. Vo, B.-N. Vo, and S. Arulampalam, "Gaussian Mixture PHD and CPHD Filtering with Partially Uniform Target Birth," in *Proceedings of the International Conference on Information Fusion*, Singapore, Jul. 2012, pp. 535–541.
- [13] 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.
- [14] —, "Extended Target Tracking using a Gaussian Mixture PHD filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3268–3286, Oct. 2012.
- [15] K. Granström and U. Orguner, "On the Reduction of Gaussian inverse Wishart mixtures," in *Proceedings of the International Conference on Information Fusion*, Singapore, Jul. 2012, pp. 2162–2169.
- [16] 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.
- [17] Y. Bar-Shalom, P. K. Willett, and X. Tian, *Tracking and data fusion, a handbook of algorithms*. YBS, 2011.
- [18] "research initiative Ko-FAS," <http://www.ko-fas.de>, 7 2012.
- [19] M. Goldhammer, E. Strigel, D. Meissner, U. Brunsman, K. Doll, and K. Dietmayer, "Cooperative multi sensor network for traffic safety applications at intersections," in *15th International IEEE Conference on Intelligent Transportation Systems*, 9 2012, pp. 1178–1183.