Extended Object Tracking assisted Adaptive Clustering for Radar in Autonomous Driving Applications

Stefan Haag*, Bharanidhar Duraisamy*, Felix Govaers#, Wolfgang Koch#, Martin Fritzsche* and Jürgen Dickmann*
(*) Deparment for Autonomous Driving, Daimler AG, Germany, Email: firstname.lastname@daimler.com,
(#) Dept. for Sensor Data and Information Fusion, Fraunhofer FKIE, Germany, Email:firstname.lastname@fkie.fraunhofer.de

***Abstract*—Multiple Extended Object Tracking in autonomous driving scenarios must be applicable in highly varying environments such as highway scenarios as well as in urban and rural environments. In this paper, a flexible UKF-based Interacting Multiple Motion (IMM) model extension for the Random Matrix Model (RMM) framework is introduced for nonlinear models. In addition to that, an adaptive clustering method where the provided tracking prior information is invoked to obtain stable clustering and tracking in varying environments with different objects and varying object types is derived. The effectiveness of the filter and clustering method is demonstrated in a real-world scenario.**

I. Introduction

The challenge of establishing autonomous mobility increases requirements for sensors, software, and algorithms, whereas two conditions seem to oppose each other. On the one hand, the utilized methods must guarantee safety for passengers of autonomous vehicles and other traffic participants. Thus, stability and robustness are essential for the realization and acceptance of automated driving. However, on the other hand, the utilized methods must provide real-time applications with restricted computational effort to make autonomous driving economically realizable. Therefore, robust and stable extended object tracking of dynamic objects in traffic scenarios is required. It is performed on fused sensor data, but in the case of sensor malfunctions, the tracking must be applied to data from a single sensor to preserve robustness. Therefore, extended object tracking on radar data only becomes essential as radar sensors are relatively robust towards external influences.

Object tracking in traffic scenarios has to be flexible as profoundly different objects and object types, among them pedestrians of all ages, vehicles in all sizes as well as all other possible road users might occur. In order to cover their motions precisely, a tracking filter has to be capable of handling varying motion models, including nonlinear models. In [1], we showed that the Random Matrix Model (RMM) is the most efficient tracking filter for radar and lidar measurements in traffic scenarios. The RMM combines the Kalman Filter for position and kinematic tracking with tracking of the spread of the obtained measurements for extent tracking. Hence, the standard RMM [2] can use only linear motion and measurement models. In this paper, an RMM extension that allows the usage of nonlinear motion and measurement models is introduced. It is based on two well-known filter, the Unscented Kalman filter (UKF) [3] and the RMM approach for nonlinear motion models [4]. Since a single motion model is not accurate enough to model the whole variety of occurring motions of all possible objects in traffic scenarios, a UKF based RMM approach for interacting multiple motion models is introduced. Besides the single model UKF-RMM base, the multiple model UKF-RMM is based on the Unscented Interacting Multiple Model filter (UIMM) [5], [6] filter, and the IMM-RMM filter [7].

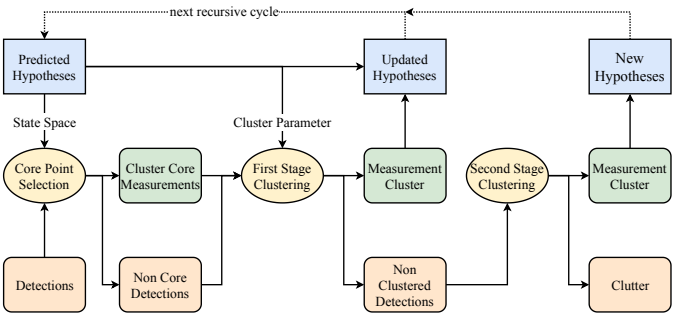


Fig. 1: Recursive cycle of the proposed extended object tracking approach (blue) and adaptive clustering method (yellow). The track hypothesis prior state are utilized to select core points from the given set of detections that form a cluster prior for each track hypothesis. With them and a track depending parameter set, the final cluster hypothesis are created with which the tracking hypotheses posterior states are calculated.They are used to compute the prior for the next cycle. Second stage clustering is performed on the set of remaining detections to find new track hypotheses

Along with that, an additional clustering step is required to make the applications real-time feasible. Clustering combines hypotheses and reduces the complexity of the measurement to track association problem significantly. Since the restriction in computational effort hold for clustering as well, many applications rely on clustering methods based on DBSCAN [8] or Optics [9] because both depend only on a few parameters. Grid-based partitions of the measurement sets allow the usage of parameters adjusted to the sensor resolution [10], [11].

Nevertheless, it occurs frequently in traffic scenarios that clustering methods with fixed-parameter sets are not suitable due to varying object sizes and characteristics.

The contribution of this paper is a UKF-based multiple motion model approach extension to the RMM extended object tracking filter for efficient and robust tracking with radar data in traffic scenarios. Furthermore, an adaptive clustering method is introduced where information is exchanged between the proposed filter and the clustering method. The information exchange of a cycle is shown in Fig 1. The blue labeled parts indicate the UIMM-RMM Filter parts. They are introduced in Section II. Adaptive clustering and its main issues, core point selection, first stage clustering, and second stage clustering, are labeled in yellow. They are derived in Section III. The obtained clusters are needed to calculate the filter posterior, and the tracking prior information is utilized to obtain cluster priors. The proposed framework is tested in Section IV. Finally, the paper is concluded with Section V.

II. Extended Object Tracking Filter

The Random Matrix Model (RMM) [12] provides tracking of length, width, and orientation along with tracking of centroid and kinematic state. The object shape is approximated with an ellipsoid that is represented as a symmetric positive definite matrix $X_k \in S^d_{++}$ in addition to the state vector $\mathbf{x}_k$ that indicates the centroid's position and dynamic state. Object measurements $Z^k = \{\mathbf{z}^j_k\}_{j=1}^{n_k},\ n_k \in \mathbb{N},\ \mathbf{z}^j_k \in \mathbb{R}^m$ are assumed to be scattered around the object's centroid with the object's extent and measurement noise $R_k \in \mathbb{S}^m_{++}$ with scale factor $z > 0$ and measurement model $H \in \mathbb{R}^{n\times m}$.

$$\mathbf{z}^j_k \sim \mathcal{N}\left(H\mathbf{x}_k, zX_k + R_k\right),\ j = 1, ..., n_k \tag{1}$$

With the measurement model $H \in \mathbb{R}^{n\times m}$. Where [13] provided a traceable factorized state density under the assumption that state vector and extent matrix are independent. With the centroid covariance matrix $P_{k|k} \in \mathbb{S}^n_{++}$, $\nu_{k|k} \in \mathbb{R}$ degrees of freedom of the Inverse Wishart distribution ($\mathcal{IW}_d$), and sum of scatter matrices $V_{k|k} \in \mathbb{S}^d_{++}$. For the set of obtained measurement up to the current time step, $\mathcal{Z}^k = \{Z^1, ..., Z^k\}$

$$\begin{aligned} p\left(\mathbf{x}_k, X_k|Z^k\right) &= p\left(\mathbf{x}_k|Z^k\right) p\left(X_k|Z^k\right) \\ = \mathcal{N}\left(\mathbf{x}_k; \mathbf{x}_{k|k}, P_{k|k}\right) &\times \mathcal{IW}\left(X_k; \nu_{k|k}, V_{k|k}\right) \end{aligned} \tag{2}$$

Although the underlying Gaussian measurement model is not entirely suitable for large vehicles, we showed that a good state approximation is possible with the RMM [12]. However, more distinguished and class-specific measurement models are required to obtain higher tracking accuracy and stability.

A. Unscented RMM Filter

In general, a nonlinear transformation of a Gaussian random variable creates a not Gaussian distributed random variable, this holds, especially in the multivariate case. Hence, covariance matrices can not be transformed directly with nonlinear transformations. For that reason, the Extended Kalman Filter uses first-order Taylor approximations. This transformation is suitable if the transformation function is well approximated by local linearization. Otherwise, the unscented transformation is more accurate as an approximation similar to the third-order Taylor is obtained [14]. Tracking in traffic scenarios requires the utilization of nonlinear measurement models and nonlinear motion models that were not intended in the original RMM framework, but the Unscented Kalman Filter (UKF) [3] can be combined with the RMM to utilize nonlinear motion and measurement models for extended object tracking. The combination of both filters is possible because both filters have an underlying Gaussian measurement model. Table I denotes

TABLE I: UKF-RMM Prediction

Input: state estimators $\mathbf{x}_{k-1|k-1}$, $P_{k-1|k-1}$, $V_{k-1|k-1}$, $\nu_{k-1|k-1}$, process noise Q_k, motion model $f(\cdot)$ and matrix motion function $M(\cdot)$

(i)	Calculate Sigma Points $\mathcal{X}_i$ and weights $W_i^{(m,c)}$ according to [3]	
(ii)	$\mathbf{x}_{k\|k-1}$	$= \sum_{i=0}^{N} W_i^{(m)} f\left(\mathcal{X}_i\right)$
(iii)	$P_{k\|k-1}$	$= Q_k$ $+ \sum_{i=0}^{N} W_i^{(c)} \left(f\left(\mathcal{X}_i\right) - \mathbf{x}_{k\|k-1}\right)\left(f\left(\mathcal{X}_i\right) - \mathbf{x}_{k\|k-1}\right)^T$
(iv)	Calculate scale factor s according to [4]	
(v)	$\nu_{k\|k-1}$	$= (d+1)$ $\cdot \left(2 + \frac{(s-d-1)(n-d-1)(\nu_{k\|k-1}-2d-2)}{sn(v-d-1)-(s-d-1)(n-d-1)(\nu_{k\|k-1}-2d-2)}\right)$
(vi)	C_2	$= \mathbb{E}\left[M(\mathbf{x}_{k\|k-1}) V_{k-1\|k-1} M(\mathbf{x}_{k\|k-1})^T\right]$
(vii)	$V_{k\|k-1}$	$= \frac{\nu_{k\|k-1}}{\nu_{k-1\|k-1}-d-1} \frac{s-d-1}{s} \frac{n-d-1}{n} C2$

As object lengths and widths are assumed to be static, $M(\cdot)$ indicates only a rotation according to the objects yaw rate

the prediction equations for the UKF-RMM filter. Analog to the UKF, sigma points and weights have to be calculated with the unscented transformation. They are required to calculate state vector and state covariance prediction for the following time step. The Extent state prediction is similar to the linear case except that the extent matrix is rotated and the degrees of freedom are adjusted.

Integrating radar Doppler measurement requires a nonlinear measurement model. To that, the update equations have to be modified as denoted in Table I. Again, sigma points and weights have to be calculated with the unscented transformation. In this case, to obtain the pseudo measurement $\hat{\mathbf{z}}_k$, innovation covariance S_k and Kalman gain K_k. With these matrices, the extent update can be calculated similarly to in the linear case. Because of the radar measurement integration, two exceptional cases have to be considered regarding the extent tracking part. At first, the measurements in polar space imply the extent scatter matrix V_k to be a covariance matrix in polar space. However, this is not desired in many cases as the objects would not appear as ellipses in cartesian space anymore. Therefore, the innovation covariance S_k, the measurements $\overline{\mathbf{z}}_k$ and $\hat{\mathbf{z}}_k$ have to be transformed into cartesian space for the extent update in step (xi) of table II, for example, with the unbiased transformation [15]or the unscented transformation. To avoid the transformation of S_k, converted measurements according to the converted measurement Kalman filter (CMKF) [15], could be used as well. Secondly, the range-rate measurements should not be incorporated in the extent tracking part. For $d < m$, the corresponding parts of S_k, $\overline{\mathbf{z}}_k$, $\hat{\mathbf{z}}_k$ and R_k have to be withdrawn after step (x) of table II.

TABLE II: UKF-RMM update

Input: state estimators $\mathbf{x}_{k|k-1}$, $P_{k|k-1}$, $V_{k|k-1}$, $\nu_{k|k-1}$, measurements $\{\mathbf{z}_k^j\}_{j=1}^{n_k}$, measurement noise R_k and measurement model $h(\cdot)$

(i) Calculate Sigma Points $\mathcal{X}_i$ and weights $W_i^{(m,c)}$ according to [3]

(ii) $\mathcal{Z}_i = h(\mathcal{X}_i)$

(iii) $\hat{\mathbf{z}}_k = \sum_{i=0}^{N} W_i^{(m)} \mathcal{Z}_i$

(iv) $\bar{\mathbf{z}}_k = \frac{1}{n_k} \sum_{j=1}^{n_k} \mathbf{z}_k^j$

(v) $Y_k = \frac{z}{\nu_{k|k-1}-2d-2} V_{k|k-1} + R_k$

(vi) $S_k = \sum_{i=0}^{N} (\mathcal{Z}_i - \hat{\mathbf{z}}_k) \cdot (\mathcal{Z}_i - \hat{\mathbf{z}}_k)^T + \frac{1}{n_k} Y_k$

(vii) $T_k = \sum_{i=0}^{N} W_i^{(c)} \left(\mathcal{X}_i - \mathbf{x}_{k|k-1}\right) \cdot (\mathcal{Z}_i - \hat{\mathbf{z}}_k)^T$

(viii) $K_k = T_k \cdot S_k^{-1}$

(ix) $\mathbf{x}_{k|k} = \mathbf{x}_{k|k-1} - K_k \cdot (\bar{\mathbf{z}}_k - \hat{\mathbf{z}}_k)$

(x) $P_{k|k} = P_{k|k-1} - K_k P_{k|k-1} K_k^T$

(xi) $\hat{N}_k = X_k^{\frac{1}{2}} S_k^{-\frac{1}{2}} (\bar{\mathbf{z}}_k - \hat{\mathbf{z}}_k) \cdot (\bar{\mathbf{z}}_k - \hat{\mathbf{z}}_k)^T S_k^{-\frac{T}{2}} X_k^{\frac{T}{2}}$

(xii) $Z_k = \sum_{j=1}^{n_k} (\mathbf{z}_k^j - \hat{\mathbf{z}}_k) \cdot (\mathbf{z}_k^j - \hat{\mathbf{z}}_k)^T$

(xiii) $\hat{Z}_k = X_k^{\frac{1}{2}} Y_k^{-\frac{1}{2}} Z_k Y_k^{-\frac{T}{2}} X_k^{\frac{T}{2}}$

(xiv) $V_{k|k} = V_{k|k-1} + \hat{N}_k + \hat{Z}_k$

Where $B = A^{\frac{1}{2}}$ indicates a square root matrix with $A = B^T B$ which is not unqiue. The matrix square roots can be calculated with Cholesky decomposition as the corresponding matrices are symmetrical positive definite.

B. Unscented Multiple Model RMM Filter

Due to the wide variety of different road users, a single motion model can not cover the whole dynamic scope of all road users. Pedestrians might be more likely to move according to the constant velocity model, whereas cars or trucks might be more likely to move according to the constant turn model. The IMM filter provides parallel tracking with multiple motion models and recognition of model switches. Thereto, it is assumed that the objects move according to one of M given motion models at a time step. Motion models switches are modeled as a Markov process with initial state probability vector $\mu_0 \in \mathbb{R}^M$ and transition probability matrix $\pi \in \mathbb{R}^{M \times M}$. Thereto, the state density (2) can rewritten with the law of total probability and the track hypotheses $\left(\mathbf{x}_k^i, P_k^i, X_k^i, \nu_k^i\right)$ for each motion model

$$\begin{aligned} p\left(\mathbf{x}_k, X_k | \mathcal{Z}^k\right) &= \sum_{j=1}^{M} p\left(\mathbf{x}_k^i, X_k^i | M_i^k, \mathcal{Z}^k\right) p\left(M_i^k | \mathcal{Z}^k\right) \\ &\approx \sum_{j=1}^{M} p\left(\mathbf{x}_k^i | M_i, \mathcal{Z}^k\right) p\left(X_k^i | M_i, \mathcal{Z}^k\right) \mu_k^j \end{aligned} \tag{3}$$

Table III shows the mixing and prediction step. First, the mixing probabilities $\mu_{k-1|k-1}^{i|j}$ are calculated in (i) and (ii) according to the transition probabilities π_{ij}. With them, the new mixed augmented state estimators $\left(x_{k-1|k-1}^{0j}, P_{k-1|k-1}^{0j}, \nu_{k-1|k-1}^{0j}, V_{k-1|k-1}^{0j}\right)$. Each of them is covering all dimensions of the complete mixed state space, including dimensions that might not be relevant for the single motion model, for example, the yaw rate for the constant velocity model. Therefore, the mixed state estimators in (viii) are obtained by merely omitting none usable dimensions of the augmented state estimators, and a prediction step analog to the previous section can be performed for each estimator set and motion model. Each individual prior is getting reinitialized in every prediction step depending on all tracking hypotheses prior weighted with probability μ_k^i

TABLE III: UIMM-RMM Mixing and Prediction

Input: state estimators $\mathbf{x}_{k-1|k-1}^i$, $P_{k-1|k-1}^i$, $V_{k-1|k-1}^i$, $\nu_{k-1|k-1}^i$, motion state $\mu_{k-1|k-1}$, transition probabilities P, process noises Q_k^i, motion models $f^i(\cdot)$ and matrix motion function $M^i(\cdot)$

(i) $\bar{c}_j = \sum_{i=1}^{M} \pi_{ij} \mu_{k-1|k-1}^i$

(ii) $\mu_{k-1|k-1}^{i|j} = \frac{1}{\bar{c}_j} \pi_{ij} \mu_{k-1|k-1}^i$

(iii) $\mathbf{x}_{k-1|k-1}^{0j} = \sum_{i=1}^{M} \mu_{k-1|k-1}^{i|j} \mathbf{x}_{k-1|k-1}^i$

(iv) $P_{k-1|k-1}^{0j} = \sum_{i=1}^{M} \mu_{k-1|k-1}^{i|j} \Big[P_{k-1|k-1}^i + \left(\left(\mathbf{x}_{k-1|k-1}^i - \mathbf{x}_{k-1|k-1}^{0j}\right) \cdot \left(\mathbf{x}_{k-1|k-1}^i - \mathbf{x}_{k-1|k-1}^{0j}\right)^T\right)\Big]$

(v) $X_k^{0j} = \sum_{i=1}^{M} \frac{\mu_{k-1|k-1}^{i|j}}{\nu_{k-1|k-1}-2d-2} V_{k-1|k-1}^i$

(vi) $\nu_{k|k-1}^{0j} = \sum_{i=1}^{M} \mu_{k-1|k-1}^{i|j} \Big[\nu_{k-1|k-1}^i + \mathrm{tr}\left(\left(X_k^i - X_k^{0j}\right)^2\right)\Big]$

(vii) $V_{k|k-1}^{0j} = \sum_{j=1}^{M} \mu_{k-1|k-1}^{j|i} \left(\nu_{k-1|k-1}^{0j} - 2d - 2\right) X_k^{0j}$

(viii) Extract $\mathbf{x}_{k-1|k-1}^i$, $P_{k-1|k-1}^i$, $V_{k-1|k-1}^i$ and $\nu_{k-1|k-1}^i$

(ix) Calculate $\mathbf{x}_{k|k-1}^i$, $P_{k|k-1}^i$, $V_{k|k-1}^i$, $\nu_{k|k-1}^i$ according to table I

TABLE IV: UIMM-RMM Update and Combination

Input: state estimators $\mathbf{x}_{k|k-1}^i$, $P_{k|k-1}^i$, $V_{k|k-1}^i$, $\nu_{k|k-1}^i$, $\mu_{k|k-1}$, measurement noise R_k and measurement model $h(\cdot)$

(i) Calcualte $\mathbf{x}_{k|k}^i$, $P_{k|k}^i$, $V_{k|k}^i$, $\nu_{k|k}^i$ according to table II

(ii) $\Lambda_k^i = \mathcal{N}\left(\bar{\mathbf{z}}_k; \hat{\mathbf{z}}_k^i, S_k^i\right)$

(iii) $\mu_{k|k}^i = \frac{\bar{c}_i \Lambda_k^i}{\sum_{j=1}^{M} \Lambda_k^i \bar{c}_j}$

(iv) $\mathbf{x}_{k|k}^i = \sum_{i=1}^{M} \mu_{k|k}^i \mathbf{x}_{k|k}^i$

(v) $P_{k|k}^i = \sum_{i=1}^{M} \mu_{k|k}^i \left[P_k^i \cdot \left(\mathbf{x}_{k|k}^i - \mathbf{x}_{k|k}\right)\left(\mathbf{x}_{k|k}^i - \mathbf{x}_{k|k}\right)^T\right]$

(vi) $V_{k|k}^i = \sum_{i=1}^{M} \mu_{k|k}^i V_{k|k}^i$

(vii) $\nu_{k|k} = \sum_{i=1}^{M} \mu_{k|k}^i \nu_{k|k}^i$

Update and combination step to obtain the combined state are shown in Table IV. An update step similar to the single motion case is performed in (i) except the calculation of weights Λ_k^i in (ii) to detect changes in the motion state. With the updated motion state from (iii), a new set of combined state estimators is calculated. Please note, that only the quality of centroid prediction is relevant to estimate which motion state is active. Step (iv) and (v) are equivalent to the point tracking case as state vector and extent matrix are assumed to be independent for all hypotheses according to(3). Analog to the linear case in [13] the extent state and the corresponding degrees of freedom are calculated with first order moment matching.

In [7], a heuristic extension to the weight calculation Λ_k^i was introduced, where the quality of the extent prediction Y_k^i was taken into account as well. The practice has shown that extent measurements are too much varying over time so that better motion state estimations are obtained without regarding the extent measurements. Nevertheless, indirectly the extent

estimator $X^i_{k|k-1}$ affects S^i_k and thereby Λ^i_k.

III. Adaptive Clustering

Clustering sensor data is a typical measurement preprocessing step to establish multiple extended object tracking with feasible complexity. For application in traffic scenarios, the clustering method must be able to handle unknown and varying numbers of clusters, measurement noise, and clutter. Stable and robust extended object tracking is not possible without consistent clustering as clustering errors provoke further tracking errors. For that reason, an adaptive clustering method that allows the usage of prior information to minimize clustering errors is introduced.

Prior knowledge considers the predicted extended tracking state. It is utilized to obtain cluster priors with a chi-squared test. Depending on the accessible information, the test can be performed with further information if a suitable metric and test can be found. For example, in [12], classification information is combined with radar position and Doppler measurements. Classification information can be further utilized to use class-specific measurement models for a better core point selection and derive models for the density of detections and the total number of detections per object. Besides, the core point selection tracking information is used in the first stage clustering step.

A. Second Stage Clustering

If no prior information is available, a second stage clustering method is required. It is needed to find untracked objects and initialize new track hypotheses. Theoretically, every clustering method that can be applied for clustering measurement data in traffic scenarios can be applied as second stage clustering, but it is recommended to use a method that can be adjusted to the sensor resolution. For the application on radar data, a grid-based approach [10] is proposed. Even though tracking information is not available, the parameters can be adjusted to the sensor resolution. Therefore, the sensor's field of view is partitioned in grid cells, where each of them is clustered separately, for example, with DBSCAN. An example of the proposed clustering on a single grid cell is shown in Fig 2. The selected grid cell covers the area next to the ego platform. Therefore, detections of an object would be rather dense, and a small ϵ value would be chosen. All dynamical measurements lying inside the grid cell and its ϵ-enclosure were used for clustering with DBSCAN. Adding the ϵ-enclosure of each grid cell ensures that overlapping clusters are found as well.

B. Core Point Selection

Core points are selected for each track separately. An acceptance region is calculated for each track separately. With the measurement equation 1 a chi-squared test is derived for single measurements $\mathbf{z}^j_k$.

$$\left(\mathbf{z}^j_k - \hat{\mathbf{z}}_k\right)\left(zX_k + R_k\right)^{-1}\left(\mathbf{z}^j_k - \hat{\mathbf{z}}_k\right) \leq \chi^2_m\left(\alpha\right) \qquad (4)$$

Each detection inside of an acceptance region is selected as a core point and assigned to the track hypothesis. An

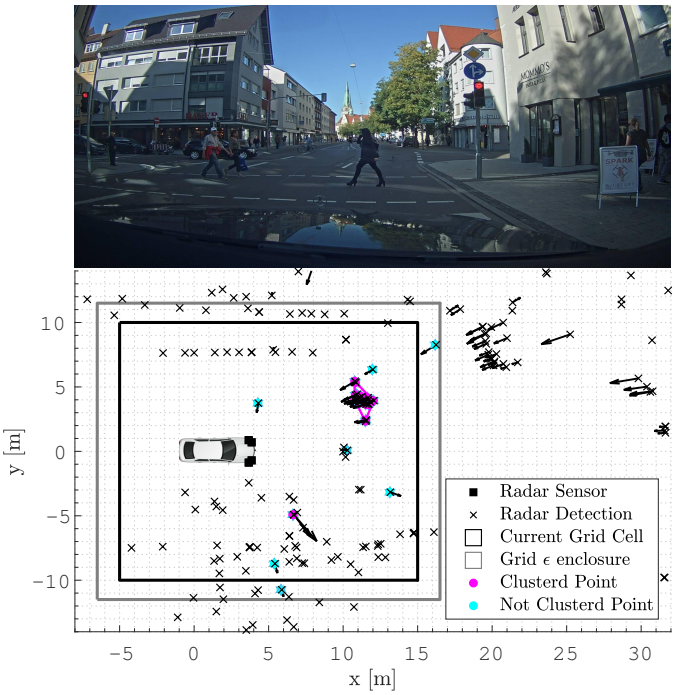


Fig. 2: Grid based clustering in an urban scenario. The scenario can be seen in the image above. Pedestrians are crossing the street in front of the halting ego vehicle. Here the grid cell around the ego vehicle is shown with its ϵ-enclosure. All detections in the enclosure with Doppler velocity over a threshold are selected for clustering. In this case, two clusters were found

example of this is shown in Fig 4. In this example the application for $d = 2$ is shown. The inclusion of range rate measurements can be realized with an augmented extension matrix $\hat{X}_k = \begin{pmatrix} X_k & 0 \\ 0 & 1 \end{pmatrix}$. Thereby, detections that are located in an acceptance region around predicted measurements are selected as core points and assigned to the track. In case a detection passes the chi-squared test for several tracks, it has to be assigned to the nearest track according to the Mahalanobis distance (4). In the case of wholly missed tracks (1) allows the sampling of core points for the first stage clustering. Independent of the distance from one point to the other and density, all detections that are assigned to the same track are considered as core points of one cluster.

C. First Stage Clustering

The objective of the first order clustering is to find one cluster out of a set of detections with given cluster core points for a given track. The clustering is performed for each track individually, which means that the clustering parameter or even the clustering method can be chosen individually for each track hypothesis. Thus, a massive variety of possibilities exist to find an optimal method and parameter coupling, where the optimal set up has to be determined for each application differently. However, this depends on the use case, for example, it has

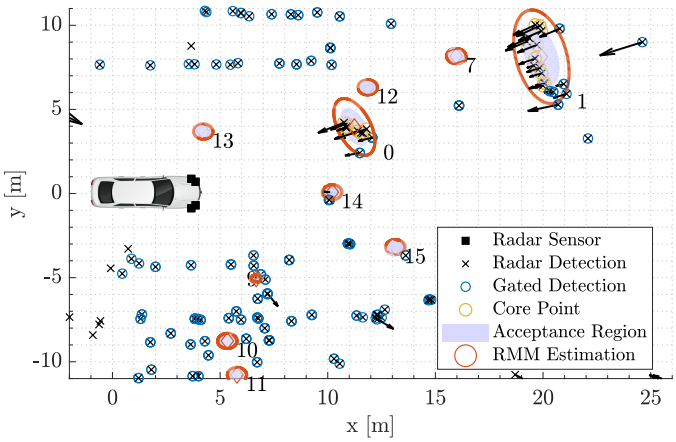


Fig. 3: Core Points are assigned to a track if they are located in the trust region of the track. The Image shows the track hypotheses as red ellipses. The acceptance regions, depending of the track position, track extent and measurement noise are shown as blue areas. Gated measurements located in the track's acceptance region are assigned as core points

to be considered if occurring objects are homogeneous or inhomogeneous, if the core points are sampled or not, if transvision or other sensor artifacts have to be considered.

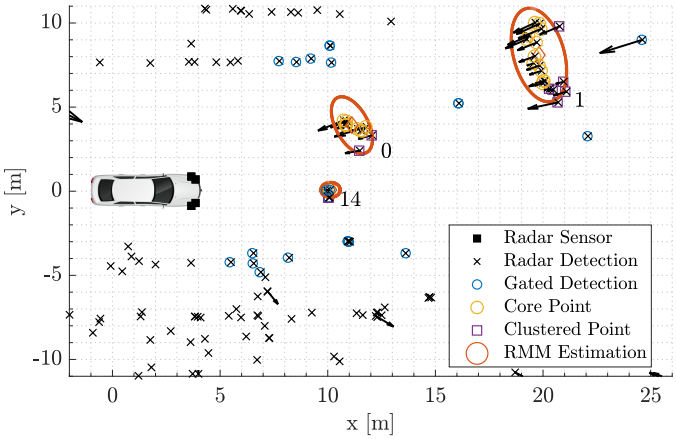


Fig. 4: Core Points are assigned to a track if they are located in the trust region of the track. The Image shows the track hypotheses as red ellipses. The acceptance regions, depending of the track position, track extent and measurement noise are shown as blue areas. Gated measurements located in the track's acceptance region are assigned as core points. Clustering can be only performed if core points exist

For the application on radar data, the clustering was performed with a DBSCAN [8] on detections $\mathbf{z}_k^j = (r_k^j, \varphi_k^j, \dot{r}_k^j)^T$ that were transformed with the unbiased transformation [16] in $\mathbb{R}^4$. The parameter set was chosen depending on the distance from sensor to object, and the Mahalanobis distance was chosen as distance metric with the track depending covariance matrix

$$D_M = \left(P_{k|k-1}\right)_{(1:4,1:4)} + \begin{pmatrix} X_{k|k-1} & 0 \\ 0 & I_{2\times 2} \end{pmatrix}. \tag{5}$$

In the case of a large number of measurements and track hypotheses, this can get computational very expensive because the measurement to track distance has to be calculated for every track hypothesis. It can be avoided by selecting candidates for each track with a gating step that excludes distant detections.

This method provides only a single clustering hypothesis per scan. Utilizing the different motion hypothesis of each track would allow multiple hypothesis clustering.

IV. Experimental Results

The clustering performance is evaluated on a manually labeled radar data set. In this data set, every dynamic object obtains a unique cluster-id. Then, labeling workers have to select all detections from the same object, and the detections receive a reference to the same cluster-id. Thereby, a set of true and false positives is obtained. True positive indicates measurements originating from real dynamic objects. True negatives can be a measurement originating from a static object or a reflection or a clutter. The cluster performance analysis is obtained by detection-wise calculating precision and recall.

$$\begin{aligned} &\text{TP: True Positive} && \text{FP: False Positive} \\ &\text{FN: False Negative} && \text{TN: True Negative} \\ &\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} && \text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \end{aligned} \tag{6}$$

Both numbers lie between zero and one where one indicates the optimal case of no false positives, respectively, no false negatives. Low precision or recall values towards zero would indicate worse clustering performance.

A. Clustering Performance Evaluation

The urban scenario from Fig 4 is used to show differences between adaptive clustering and clustering without prior information. The clustering results of 5 consecutive steps are shown in Fig 5. Table VI shows the corresponding precision and recall values. In general, it is more likely to obtain higher recall values than precision values because false positives appear regularly as a sensor-specific artifact. Nevertheless, the

TABLE V: Step-wise Clustering Performance Comparison

Step	1	2	3	4	5
Precision AC1	0.77	0.78	0.81	0.81	0.79
Precision AC3	-	-	0.97	0.95	0.95
Precision GBC	0.77	0.76	0.76	0.60	0.85
Recall AC 1	0.92	0.93	0.97	0.96	0.92
Recall AC 3	-	-	0.78	0.83	0.88
Recall GBC	0.92	0.93	0.84	0.69	0.95

Precision and recall values for adaptive clustering evaluated with all track hypotheses (AC1), adaptive clustering performance of tentative tracks that were observed more than 2 cycles (AC3) and grid-based DBSCAN adaptive clustering performance (GBC) for each step of the scenario shown in Fig 5

precision is improved if only tracks are used that were observed for more than two measurement cycles (AC3), but this requires a longer initialization time. Adaptive clustering and grid-based clustering provide recall values close to one, which indicates that missed detections are significantly less frequent than false positives. Grid-based clustering may outperform adaptive clustering in single steps, for example, in step five, but adaptive clustering denotes stable clustering results concerning the recall values if accurate prior information is available. Step four and five show that grid-based clustering is more likely to have outlying clustering results. However, clustering errors can be propagated with adaptive clustering. The long term

TABLE VI: Average Clustering Performance Comparison

	urban scenario	highway scenario
Precision AC1	0.64	0.13
Precision AC3	0.74	0.22
Precision AC9	0.64	0.25
Precision GBC	0.70	0.76
Recall AC 1	0.82	0.70
Recall AC 3	0.79	0.69
Recall AC 9	0.79	0.72
Recall GBC	0.80	0.63

Average precision and recall values for adaptive clustering evaluated with all track hypotheses (AC1), adaptive clustering performance of tentative tracks that were observed more than 2 cycles (AC3), verified tracks (AC9) and grid-based DBSCAN adaptive clustering performance (GBC) for the complete urban scenario and a highway scenario.

observations, shown in Table VI that the tendency towards a higher number of false positives becomes more apparent, especially in the highway scenario. This may be caused by an augmented number of reflections caused by metal surfaces such as road boundaries and other cars. Nevertheless, $80\,\%$ of all measurements in the urban scenario and $70\,\%$ of all measurements in the highway scenario are clustered and with adaptive clustering automatically tracked as well over several measurement cycles. The grid-based clustering provides similar recall values but only for a single measurement cycle.

V. Conclusion

The standard RMM has been limited to linear measurement and motion models. This paper provides an unscented interacting multiple model RMM extension for efficient and robust tracking in different automotive scenarios. To obtain stability, a new clustering method that allows the inclusion of tracking prior information into the clustering process. We showed that with the combination of the multiple model extended object filter and the adaptive clustering and mediately tracking of extended objects with highly varying sizes, characteristics, and dynamic behaviors without remarkable extra computational effort. The new framework allows better clustering and, thereby, better and more stable tracking in noisy environments unnecessary. The framework can be adapted to different scenarios, as all parts are interchangeable. To avoid error propagation by wrong prior information, an optimal adjustment, considering both the tracking filter and the clustering, is required for every transfer to a different application or different sensor set up.

Acknowledgment

The authors would like to thank Dr. Hans-Ludwig Blöcher of Team Radar for the valuable discussions. A part of this work is funded by German Federal Ministry of Education and Research (BMBF) and Federal Ministry for Economic Affairs and Energy (BMWi) through Radar4FAD project with the project grant no. 16ES0560.

References

[1] S. Haag, B. Duraisamy, W. Koch, and J. Dickmann, "Radar and Lidar Target Signatures of Various Object Types and Evaluation of Extended Object Tracking Methods for Autonomous Driving Applications," in *018 21st International Conference on Information Fusion (FUSION)*, (Cambridge), pp. 1746–1755, IEEE, 2018.

[2] W. Koch, "On Bayesian tracking of extended objects," *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 209–216, 2006.

[3] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," *IEEE Transactions on Automatic Control*, vol. 45, no. 2, pp. 477–482, 2000.

[4] K. Granstrom and U. Orguner, "New prediction for extended targets with random matrices," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 1577–1589, 2014.

[5] M. Djouadi, A. Sebbagh, and D. Berkani, "IMM-UKF algorithm and IMM-EKF algorithm for tracking highly maneuverable target: a comparison," *Proceedings of the 7th WSEAS International Conference on Automatic Control, Modeling and Simulation*, pp. 283–288, 2005.

[6] J. Hartikainen, A. Solin, and S. Särkkä, "Optimal Filtering with Kalman Filters and Smoothers a Manual for the Matlab toolbox EKF/UKF," *Department of Biomedical Engineering and Computational Science, Aalto University School of Science, P.O.Box 1100, FI-00076 AALTO, Espoo, Finland*, 2011.

[7] M. Feldmann and D. Fränken, "Advances on Tracking of Extended Objects and Group Targets using Random Matrices," *2009 12th International Conference on Information Fusion*, pp. 1029–1036, 2009.

[8] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, "A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise," *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, pp. 226–231, 1996.

[9] M. Ankerst, M. M. Breunig, H.-p. Kriegel, and J. Sander, "OPTICS: Ordering Points To Identify the Clustering Structure," in *Proceedings of the 1999 ACM SIGMOD international conference on Management of data - SIGMOD '99*, no. June, (New York, New York, USA), pp. 49–60, ACM Press, 1999.

[10] S. Mahran and K. Mahar, "Using grid for accelerating density-based clustering," *Proceedings - 2008 IEEE 8th International Conference on Computer and Information Technology, CIT 2008*, pp. 35–40, 2008.

[11] D. Kellner, J. Klappstein, and K. Dietmayer, "Grid-based DBSCAN for clustering extended objects in radar data," *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 365–370, 2012.

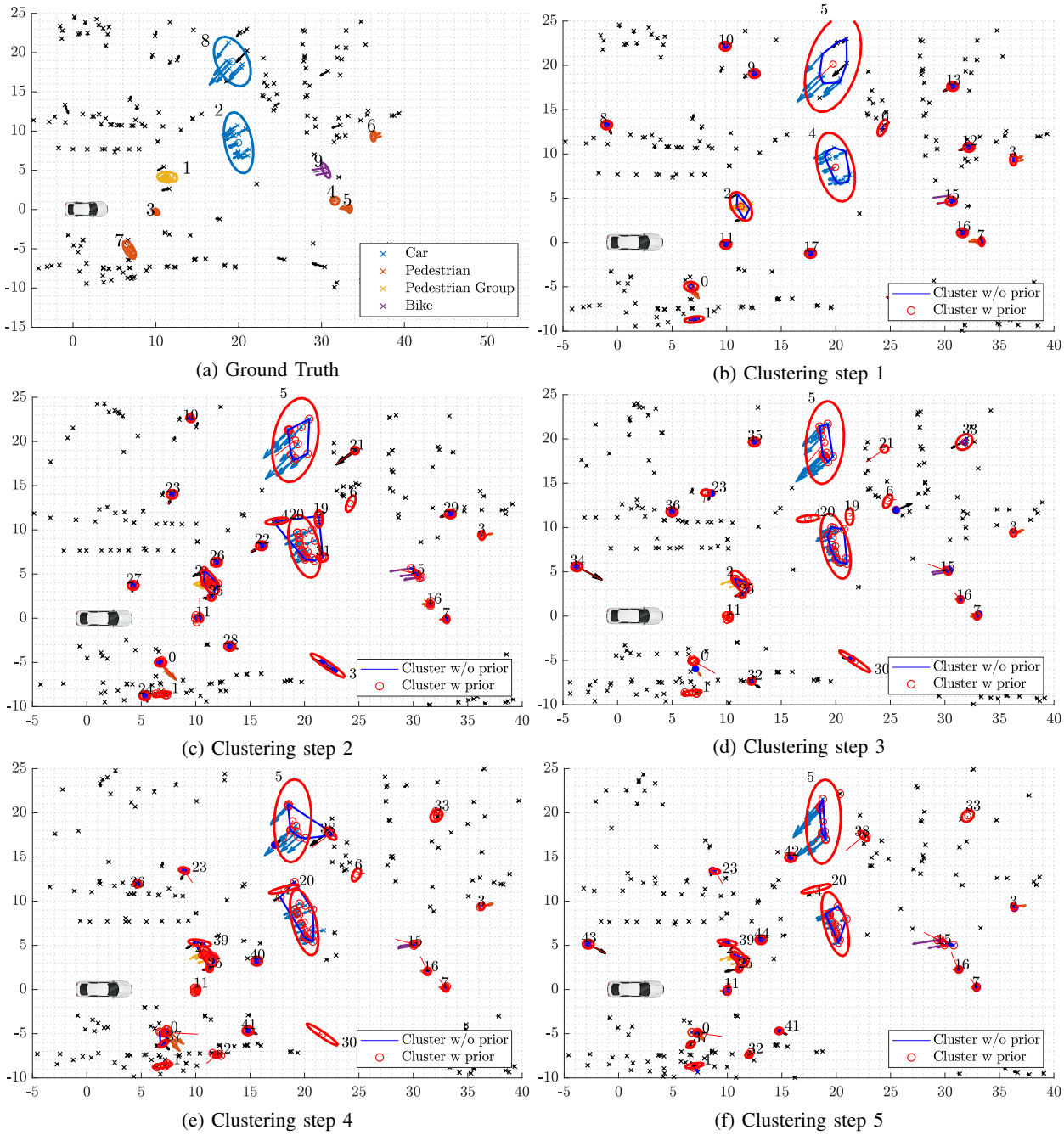
[12] S. Haag, B. Duraisamy, W. Koch, and J. Dickmann, "Classification Assisted Tracking for Autonomous Driving Domain," in *2018 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, (Bonn), pp. 1–8, 2018.

[13] M. Feldmann, D. Fränken, and W. Koch, "Tracking of extended objects and group targets using random matrices," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1409–1420, 2011.

[14] E. A. Wan and R. van der Merwe, "The Unscented Kalman Filter for Nonlinear Estimation," *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*, pp. 153–158, 2000.

[15] S. V. Bordonaro, P. Willett, and Y. Bar-Shalom, "Unbiased tracking with converted measurements," in *IEEE National Radar Conference - Proceedings*, pp. 0741–0745, 2012.

[16] S. V. Bordonaro, P. Willett, and Y. Bar-Shalom, "Consistent linear tracker with position and range rate measurements," *Conference Record - Asilomar Conference on Signals, Systems and Computers*, pp. 880–884, 2012.



(a) Ground Truth

(b) Clustering step 1

(c) Clustering step 2

(d) Clustering step 3

(e) Clustering step 4

(f) Clustering step 5

Fig. 5: Clustering results of an urban with adaptive clustering in red in comparison to DBSCAN clustering. DBSCAN is performed with position and velocity information obtained from unbiased transformed [15] radar measurements. The blues lines indicate the convex hull of a cluster. Detections obtained from adaptive clustering are labeled as red circles with the corresponding estimated ellipse. Fig 5a shows the ground truth objects in the initial step with a pedestrian group (1) crossing the street in front of the ego vehicle, four pedestrians (4,5,6,7) crossing the street, two cars (2,8) approaching the intersection and a cyclist (9) performing a right turn, this also shown in Fig 2. The following Fig 5b to Fig 5f show the clustering results. The cluster results in step 1 are identical as no prior exists but differentiate in the following steps.