MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

Extended Object Tracking with Automotive Radar Using B-Spline Chained Ellipses Model

Yao, Gang; Wang, Pu; Berntorp, Karl; Mansour, Hassan; Boufounos, Petros T.; Orlik, Philip V.

TR2021-048 May 11, 2021

Abstract

This paper introduces a B-spline chained ellipses model representation for extended object tracking (EOT) using high-resolution automotive radar measurements. With offline automotive radar training datasets, the proposed model parameters are learned using the expectationmaximization (EM) algorithm. Then the probabilistic multi-hypothesis tracking (PMHT) along with the unscented transform (UT) is proposed to deal with the nonlinear forward-warping coordinate transformation, the measurement-to-ellipsis association, and the state update step. Numerical validation is provided to verify the effectiveness of the proposed EOT framework with automotive radar measurements.

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

EXTENDED OBJECT TRACKING WITH AUTOMOTIVE RADAR USING B-SPLINE CHAINED ELLIPSES MODEL

G. Yao, P. Wang, K. Berntorp, H. Mansour, P. Boufounos, and P. V. Orlik*

Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA

ABSTRACT

This paper introduces a B-spline chained ellipses model representation for extended object tracking (EOT) using high-resolution automotive radar measurements. With offline automotive radar training datasets, the proposed model parameters are learned using the expectation-maximization (EM) algorithm. Then the probabilistic multi-hypothesis tracking (PMHT) along with the unscented transform (UT) is proposed to deal with the nonlinear forward-warping coordinate transformation, the measurement-to-ellipsis association, and the state update step. Numerical validation is provided to verify the effectiveness of the proposed EOT framework with automotive radar measurements.

***Index Terms*—** Extended object tracking, automotive radar, unscented transform, PMHT, B-spline, random matrix model.

1. INTRODUCTION

With increasingly higher angular resolution and rapid advances in automotive radar, more and more detection points per time scan are obtained for a single object and, as a result, extended object tracking (EOT) is well suited to summarize the statistics from the multiple detection points and track the object. Compared with traditional point object tracking, EOT can estimate not only the kinematic state but also the extent state including the length and width of objects [1].

One key issue in EOT is to capture the spatial representation of multiple detection points given the object state including the position, orientation, length, and width. EOT is commonly used together with LIDAR measurements and/or image processing than automotive radar. As a consequence, the spatial representation can be classified into two main categories: 1) contour models including rectangular shape models [2, 3], the star-convex shape models [4–8], and B-spline curve [5], which reflect the measurement distribution along the object contour, a case suitable to the LIDAR application; and 2) surface models such as the random matrix model [9–12], image moments [13] and a graph model [14], which assume that the measurements are generated from the object surface.

Real-world automotive radar measurements are, however, more complex and can neither be well described by the contour model nor by the surface model, see, e.g., [15, 16]; see Fig. 1 for an illustration of the real-world accumulated automotive radar measurements [17] in a unit coordinate where the origin is located at the middle of the rear axle. This has motivated the third category of models, the surface-volume models, to balance between the contour models and the surface models with more realistic features customized to the automotive radar measurements. Typical examples of surface-volume models in the literature include the volcanormal measurement model [18], the variational Gaussian mixture model (GMM) [17], and the hierarchical truncated Gaussian (HTG) model [19–22].

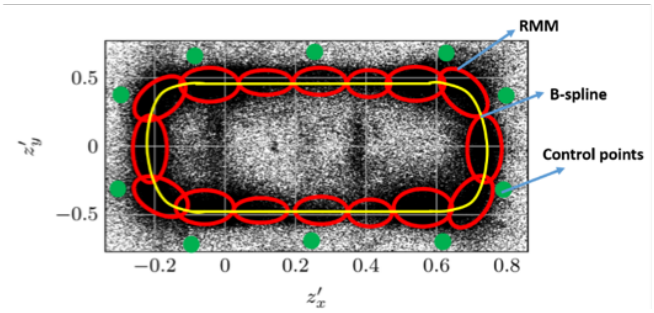


Fig. 1. Accumulated automotive radar measurements in a unit coordinate [17] and a B-spline chained ellipses model.

In this paper, we combine the contour and surface models to introduce a new surface-volume model for the automotive radar measurements. As illustrated in Fig. 1, we propose a regularized multiple ellipses model to describe the spatial representation of automotive radar measurements given the object state. Such a regularization is enforced by the requirement that the center of each ellipsis component (e.g., a typical surface model) has to be located on an enclosed B-spline curve that represents the vehicle contour (e.g., a typical contour model). Each ellipsis component is used to cover the measurement spread along the B-spline object contour. Our contour-based regularization of multiple ellipses is different from the variational Gaussian mixture model of [17], where a hierarchical signal model with prior distributions on ellipsis means and covariance matrices, and a prior Dirichlet distribution on the mixture coefficient, is used to regularize the spatial representation and the conventional Gaussian mixture model of [23], where no regularization was introduced.

With the proposed spatial representation model, we propose to learn associated model parameters from offline automotive radar measurements via the expectation-maximization (EM) algorithm [24]. Once the proposed B-spline chained ellipses model is learned, the unscented Kalman filter (UKF) [25] is used for forward-warping the learned measurement model from the unit coordinate into the global coordinate of the automotive radar measurements. Probabilistic multi-hypothesis tracking (PMHT) is then applied to assign the automotive measurements to different ellipsis components in a probabilistic fashion and update both the kinematic and extent states.

*The work of G. Yao was done during his internship at MERL.

2. LEARNED B-SPLINE CHAINED ELLIPSES MODEL

2.1. Proposed B-Spline Chained Ellipses Model

The proposed spatial model consists of L Gaussian components (i.e., ellipses) with their component means located on a B-spline curve. For each ellipse centered at μ_l with an extent Σ_l, we can associate the N_k measurements with an association probability ρ_i^l. Given the measurement-to-ellipse assignment, the likelihood function becomes

$$\phi\left(\mathbf{Z}_k|N_l,\mu_l,\Sigma_l,\rho^l\right)\propto\mathcal{N}\left(\bar{\mathbf{z}}_l;\mu_l,\frac{\sum_l}{\bar{N}_l}\right)\times W\left(\overline{Z}_l;\bar{N}_l-1,\Sigma_l\right)$$

where $\bar{N}_l=\sum_{i=1}^{N_k}\rho_i^l$,

$$\bar{\mathbf{z}}_l=\frac{\sum_{i=1}^{N_k}\rho_i^l\tilde{\mathbf{z}}_i}{\sum_{i=1}^{N_k}\rho_i^l}, \tag{1}$$

$$\overline{\mathbf{Z}}_l=\sum_{i=1}^{N_k}\rho_i^l\left(\tilde{\mathbf{z}}_i-\bar{\mathbf{z}}_l\right)\left(\tilde{\mathbf{z}}_i-\bar{\mathbf{z}}_l\right)^{\mathrm{T}}, \tag{2}$$

are, respectively, the sample mean and spread of the l-th ellipse, $\mathcal{N}$ denotes the Gaussian distribution and $\mathcal{W}$ is the Wishart distribution. With all L ellipses and given the measurement-to-ellipse association, the L random matrices model is defined as

$$p(\mathbf{Z}|\theta,\rho)=\sum_{l=1}^{L}\pi_l\phi\left(\mathbf{Z}_k|N_l,\mu_l,\Sigma_l,\rho\right), \tag{3}$$

where the mixture weights π_l are assumed to equal $\pi_l=1/L$.

Moreover, we assume that the ellipse centers are located on a B-spline curve define by $\mathbf{c}(r)\in\mathbb{R}^{2\times1}$ of degree d [26]

$$\mathbf{c}(r)=\sum_{j=0}^{m}\mathbf{p}_jB_{j,d}(r),\quad 0\le r\le m-d+1, \tag{4}$$

where $\mathbf{p}_j\in\mathbb{R}^{2\times1}$ is the j-th control point, $m+1$ is the number of control points, and $B_{j,d}(r)$ is the basis function with a parameter r [26]. By enforcing $\mu_l=\mathbf{c}(r_l)$ with r_l denoting the corresponding parameter of the l-th ellipse center μ_l, the B-spline chained ellipses model is defined as

$$p(\mathbf{Z}|\theta,\rho)=\sum_{l=1}^{L}\pi_l\phi\left(\mathbf{Z}_l|N_l,\mathbf{c}(r_l),\Sigma_l,\rho\right), \tag{5}$$

where the parameters of the proposed model are the number of measurements for each component N_l, control points of the B-spline curve $\{\mathbf{p}_j\}_{j=0}^{m}$ and the covariance matrices of each component $\{\Sigma_l\}_{l=1}^{L}$.

2.2. Offline Model Learning via EM

To learn the model parameter in the above section, we apply the following coordinate transformation to convert the training dataset $\mathbf{Z}=\{\mathbf{z}_i\}_{i=1}^{N}$ into a unit coordinate system that originates at the object center $\mathbf{m}=[x_m,y_m]\in\mathbb{R}^{2\times1}$ and is oriented such that the x-axis points towards the front:

$$\tilde{\mathbf{z}}_i=S^{-1}R_\psi^{-1}\left(\mathbf{z}_i-\mathbf{m}\right), \tag{6}$$

where $R_\psi\in\mathbb{R}^{2\times2}$ is the rotation matrix as a function of the orientation angle ψ and $S=\mathrm{diag}(l,w)$ is a scaling matrix. With the transformed training measurements $\tilde{\mathbf{Z}}=\{\tilde{\mathbf{z}}_i\}_{i=1}^{N}$, we derive the EM algorithm to estimate the associated parameters (i.e., hidden random variables and deterministic model parameters) of the B-spline chained ellipses model.

Expectation step is to update the hidden random variables $\{\rho_l,\bar{\mathbf{z}}_l,\overline{\mathbf{Z}}_l\}$. First we calculate the mixture weights from the posterior association probability ρ_i^l for $\tilde{\mathbf{z}}_i$ [27]

$$\rho_i^l=\frac{\frac{1}{L}\times\mathcal{N}\left(\tilde{\mathbf{z}}_i;\mu_l,4\Sigma_l\right)}{\frac{1}{L}\times\sum_{l=1}^{L}\mathcal{N}\left(\tilde{\mathbf{z}}_{i,};\mu_l,4\Sigma_l\right)+\lambda}, \tag{7}$$

where μ_l and $4\Sigma_l$ are, respectively, the mean and covariance matrix of each component, the scaling factor 4 is used to approximate a uniform distribution [10] and λ is the probability of the uniformly distributed outliers. Then, the remaining hidden variables $\bar{\mathbf{z}}_l$ and $\overline{\mathbf{Z}}_l$ can be updated using (1) and (2), respectively.

Maximization step is to update the model parameters $\theta=\{\mathbf{p}_j,\Sigma_l\}$ based on the Q-function of (5) as

$$\mathcal{Q}(\theta)\propto\sum_{l=1}^{L}\left\{-\frac{\bar{N}_l}{2}\left(\mu_l-\bar{\mathbf{z}}_l\right)^{\mathrm{T}}\Sigma_l^{-1}\left(\mu_l-\bar{\mathbf{z}}_l\right)-\frac{\bar{N}_l+1}{2}\log|\Sigma_l|-\frac{1}{2}\mathrm{tr}\left(-\frac{1}{2}\overline{\mathbf{Z}}_l\Sigma_l^{-1}\right)\right\}. \tag{8}$$

We can reformat the B-spline curve in a matrix-vector form as $\mu_l=\mathbf{B}_l\mathbf{p}$, where $\mathbf{B}_l=\mathrm{blkdiag}\left(\mathbf{n}_l^{\mathrm{T}},\mathbf{n}_l^{\mathrm{T}}\right)$, $\mathbf{n}_l=\left[B_{0,d}\left(r_l\right),\cdots B_{m,d}\left(r_l\right)\right]^{\mathrm{T}}$, and $\mathbf{p}=\left[\mathbf{p}_x^{\mathrm{T}},\mathbf{p}_y^{\mathrm{T}}\right]^{\mathrm{T}}$ with $\mathbf{p}_x^{\mathrm{T}}$ and $\mathbf{p}_y^{\mathrm{T}}$ denoting the control points in the x- and y- coordinates, respectively. By setting the derivative of $\mathcal{Q}(\theta)$ (with respect to θ) to 0, we have $\mathbf{p}=\mathbf{H}^{+}\mathbf{M}$, where $\mathbf{H}^{+}$ is the Moore-Penrose inverse of $\mathbf{H}=\sum_{l=1}^{L}\left(\bar{N}_l\mathbf{B}_l^{\mathrm{T}}\Sigma_l^{-1}\mathbf{B}_l\right)$ and $\mathbf{M}=\sum_{l=1}^{L}\left(\bar{N}_l\mathbf{B}_l^{\mathrm{T}}\Sigma_l^{-1}\overline{\mathbf{z}}_l\right)$, and

$$\Sigma_l=\frac{1}{\bar{N}_l+1}\left[\bar{N}_l\left(\bar{\mathbf{z}}_l-\mu_l\right)\left(\bar{\mathbf{z}}_l-\mu_l\right)^{\mathrm{T}}+\overline{\mathbf{Z}}_l^{\mathrm{T}}\right]. \tag{9}$$

We then iterate between the estimates of $\mathbf{p}$ and Σ_l until the convergence is achieved, where the convergence criterion can be a predetermined likelihood in (8), the relative changes of the estimated parameters over consecutive iterations, or a predetermined maximum iteration number.

3. ONLINE TRACKING USING PMHT AND UKF

This section introduces the UKF-PMHT tracking algorithm with the offline learned spatial model. In a nutshell, we first use the UKF to compute the approximate mean and covariance of ellipse-assigned measurements in the global coordinate system, which are nonlinearly converted from the predicted kinematic and extent states and the corresponding measurement in the unit coordinate system. Then, the PMHT is used to associate the measurements to the ellipses in the global coordinates and update the object state. Compared with the local linearization-based extended Kalman filter and computationally more expensive particle filter, the UKF directly approximates the posterior by a Gaussian density represented by a set of deterministically chosen sample points and, hence, achieves a trade-off between the state estimation accuracy (over the EKF) and the complexity (over the particle filter) [25].

3.1. Forward-Warping Unscented Transformation

Given the learned B-spline chained ellipses model and assuming a measurement $\mathbf{x}_\mu$ in the unit coordinate system is distributed

with respect to the l-th ellipse $\mathcal{N}(\mu_l, \Sigma_l)$, the corresponding measurement $h_{l,k}(\mathbf{x}_{k|k-1})$ in the global coordinate system is defined as

$$h_{l,k}(\mathbf{x}_{k|k-1}) = \mathbf{m}_{k|k-1} + R_{\psi_{k|k-1}} \cdot S_{k|k-1} \cdot \mathbf{x}_\mu \tag{10}$$

where $\mathbf{m}_{\psi_{k|k-1}}$, $R_{\psi_{k|k-1}}$ and $S = \text{diag}(l_{k|k-1}, w_{k|k-1})$ are defined the same way as (6) except that all augments are given by the predicted state $(k|k-1)$ with corresponding predictive distributions (e.g., the Gaussian distribution).

Since the above transformation in (10) is nonlinear, particularly with respect to the predictive orientation angle, we apply the unscented transform (UT) [25] to determine the mean $\bar{h}_{l,k}(\mathbf{x}_{k|k-1})$ and covariance matrix $\mathbf{X}_l$ of $h_{l,k}(\mathbf{x}_{k|k-1})$. To this end, we augment the predicted state with $\mathbf{x}_\mu$ as $\mathbf{x}_{aug} = \left[\mathbf{x}_{k|k-1}, \mathbf{x}_\mu\right]^{\mathrm{T}} \in \mathbb{R}^{n_a \times 1}$ with $n_a = 9$. Then, $2n_a + 1$ weighted sample points $(\mathcal{A}_i, \mathcal{W}_i)$, i.e., the sigma points, are determined such that they can describe the true mean $\bar{\mathbf{x}}_{aug}$ and covariance matrix $\mathbf{P}_{aug}$ of $\mathbf{x}_{aug}$:

$$\mathcal{A}_0 = \bar{\mathbf{x}}_{aug}, \quad \mathcal{W}_0 = \kappa/(n_a + \kappa), \quad \mathcal{W}_{i\geq 1} = 0.5(n_a + \kappa),$$

$$\mathcal{A}_{i\leq n_a} = \bar{\mathbf{x}}_{aug} + \left(\sqrt{(n_a + \kappa)\mathbf{P}_{aug}}\right)_i, \tag{11}$$

$$\mathcal{A}_{i>n_a} = \bar{\mathbf{x}}_{aug} - \left(\sqrt{(n_a + \kappa)\mathbf{P}_{aug}}\right)_{i-n_a}, \tag{12}$$

where κ is a scaling parameter such that $\kappa + n_a \neq 0$ and $(\sqrt{\mathbf{A}})_i$ denotes the i-th row of the matrix square root of $\mathbf{A}$. Each sigma point is then propagated through the nonlinear function of (10), i.e., $\mathcal{B}_i = h_{l,k}(\mathcal{A}_i)$, and the first two moments of $h_{l,k}(\mathbf{x}_{k|k-1})$ are computed as

$$\bar{h}_{l,k} = \sum_{i=0}^{2n_a} \mathcal{W}_i \mathcal{B}_i, \tag{13}$$

$$\mathbf{X}_l = \sum_{i=0}^{n_a} \mathcal{W}_i (\mathcal{B}_i - \bar{h}_{l,k})(\mathcal{B}_i - \bar{h}_{l,k})^T. \tag{14}$$

In the global coordinate system, a measurement $\mathbf{z}_i$ that is assigned to the l-th ellipse can be defined as

$$\mathbf{z}_i = h_{l,k}(\mathbf{x}_{k|k-1}) + \mathbf{n}_l, \tag{15}$$

where $h_{l,k}(\mathbf{x}_{k|k-1}) \sim \mathcal{N}(\bar{h}_{l,k}, \mathbf{X}_l)$ is the corresponding reflection center, and $\mathbf{n}_l \sim \mathcal{N}(0, \mathbf{R})$ is the measurement noise.

Remark: It is worth noting differences between our approach and the one in [23]. First, as we stated earlier, we impose a regularization on the mixture model via the B-spline contour, while [23] used a standard Gaussian mixture model. This results in a different Maximization step in the previous section where our focus is to maximize the likelihood function with respect to the B-spline associated parameters. Second, with the learned spatial model in the unit coordinate, we use the UT to convert the learned likelihood function from the unit coordinate system to the global coordinate system, which is referred to as the forward-warping likelihood transformation. On the other hand, [23] invokes an implicit measurement model transforming new measurements in the global frame into the local frame as a function of the unknown state vector at time k, which can be referred to as the backward-warping likelihood transformation.

3.2. Probabilistic Multi-Hypothesis Tracking

Given $\mathbf{Z}_k = \{\mathbf{z}_{i,k}\}_{i=1}^{N_k}$ and L offline learned $\{h_{l,k}(\mathbf{x}_{k|k-1})\}_{l=1}^{L}$ obtained at time step k, the PMHT is used to assign the measurements to each ellipsis component. Different from the general PMHT algorithm [27] to handle the *measurement-to-object association* and update the *kinematic* states of *multiple* objects over *consecutive* time steps, the PMHT algorithm here is applied to handle the *measurement-to-ellipsis association* and update both *kinematic and extent* states of a *single* object over the *current* time step. Specifically, the PMHT employs the EM algorithm for a soft measurement-to-ellipsis assignment that in turns creates a synthetic measurement for each component. Mathematically, the measurement-to-ellipsis association weights $\rho_{i,k}^l$, synthetic measurements $\bar{\mathbf{z}}_{l,k}$, and corresponding synthetic covariance matrix $C_{\mathbf{zz}}$ are derived as follows[1]

$$\rho_{i,k}^l = \frac{\mathcal{N}\left(\mathbf{z}_{i,k}; \bar{h}_{l,k}(\mathbf{x}_{k|k-1}), 4\mathbf{X}_{l,k} + \mathbf{R}\right)}{\sum_{l=1}^{L_k} \mathcal{N}\left(\mathbf{z}_{i,k}; \bar{h}_{l,k}(\mathbf{x}_{k|k-1}), 4\mathbf{X}_{l,k} + \mathbf{R}\right)}, \tag{16}$$

$$\bar{\mathbf{z}}_{l,k} = \frac{\sum_{i=1}^{N_k} \rho_{i,k}^l \mathbf{z}_{i,k}}{\sum_{i=1}^{N_k} \rho_{i,k}^l}, \tag{17}$$

$$C_{\mathbf{zz}} = 4\mathbf{X}_{l,k} + \frac{\mathbf{R}}{\sum_{i=1}^{N_k} \rho_{i,k}^l}. \tag{18}$$

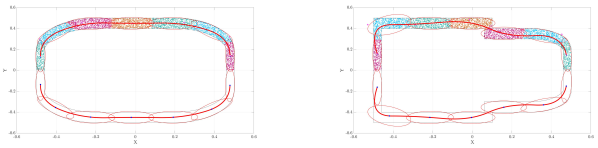
Then, the covariance between the states and measurements $C_{\mathbf{xz}}$ is calculated during the UT procedure in (10) and the filter gain is calculated as $K = C_{\mathbf{xz}} C_{\mathbf{zz}}^{-1}$. The object states $\mathbf{x}_{k,l}$ and covariance $C_{k,l}$ are updated based on the l-th measurement equation in (15), i.e., the update stage of UKF.

The PMHT iterates between the expectation and maximization steps until the predefined maximum iteration number N_{iter} is reached. In each iteration n, the object state $\mathbf{x}_{l,n}$ and covariance $C_{l,n}$ are updated incrementally by each component (i.e., over l) in the order of (10) and (16)-(18). The initial states and covariance are $\mathbf{x}_{l=0,n=1} = \mathbf{x}_{k|k-1}$ and $C_{l=0,n=1} = C_{k|k-1}$. The overall online tracking algorithm is shown in Algorithm 1.

Algorithm 1: The UKF-PMHT Tracking algorithm .

```
Set the maximum iteration number N_iter;
Set the initial state vector x_0 and covariance C_0 ;
Obtain offline learned mean {μ_l}_1^L and spread {Σ_l}_1^L ;
while Tracking do
    State x_{k|k-1} and Covariance C_{k|k-1} are predicted
      using UT and CT with polar velocity;
    Obtain measurements Z_k = {z_{k,l}}_{l=1}^{L_k} at time k;
    n= 1;x_{l=0,n=1} = x_{k|k-1};C_{l=0,n=1} = C_{k|k-1};
    while n < N_iter and not converged do
        for l = 1 to L do
            for f = 1 to L do
                Calculate the mean h̄_{f,k}(x_{l-1,n}) and
                  covariance X_{f,k} using UT and (10);
            Calculate weights ρ_{i,k}^l using (16);
            Calculate the synthetic measurement z̄_{l,k} using
              (17) and covariance C_zz using (18);
            Calculate filter gain K = C_xz C_zz^{-1};
            Update state x_{l,n} and covariance C_{l,n};
        n = n + 1;
    x_{k|k} = x_{L,N_iter}, C_{k|k} = C_{L,N_iter}
```

[1] Due to space limit, we skip the detailed derivation of the PMHT and use Algorithm 1 for the overall state update step.



(a) (b)

Fig. 2. Offline learned B-spline chained ellipses models for (a) a sedan and (b) a truck from offline measurements in the unit coordinate system.

4. SIMULATION RESULTS

In this section, we consider two vehicle types: a) a sedan and b) a truck as shown in Fig. 2 to verify the offline EM learning step and evaluate the online tracking performance. The length and width of both vehicles are 5m and 2m, respectively.

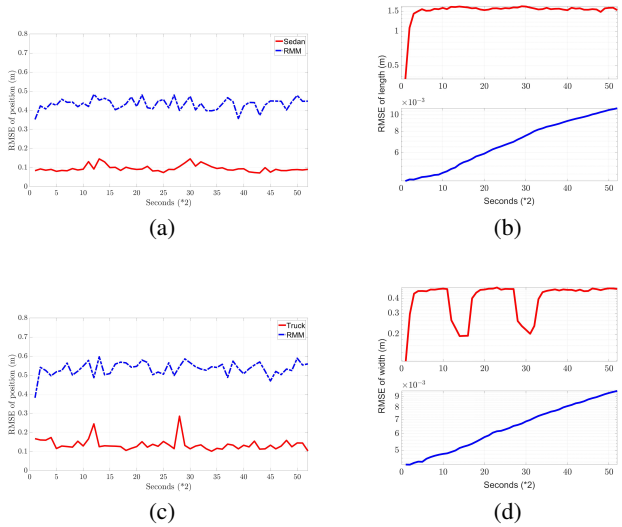
4.1. Offline Model Learning

To verify the offline EM learning step, we synthesize automotive radar measurements according to a uniform distribution within the range of 0.2m inside the true vehicle contour and then convert these offline measurements to the unit coordinate system according to (6) with labeled object state. Due to the symmetry, the model is first trained using only measurements in the positive y-coordinate region and then flipped with respect to the x-axis to represent the other half. Particularly, we set $d = 2$, $m + 1 = 6$ and $L = 7$ in (4) for the B-spline.

Fig. 2 shows the learned B-spline chained ellipses model for the two considered vehicle types in the unit coordinate system. Note that the proposed model is capable of capturing both the object contour and the measurement spread for different vehicles. For the sedan type, the offline learned model demonstrates a good balance between the measurement spread and the smoothness of the chained ellipses over the B-spline contour. On the other hand, the offline learned model for the truck shows larger discrepancy over the sharp corners around the truck. Nevertheless, the learned B-spline contour can still sufficiently reflect those sharp contour variations, while corresponding ellipsis components adapt to those sharp variations by increasing their extents (eigenvalues of the covariance matrices) and tilting their orientation angles (eigenvectors of the covariance matrices).

4.2. Online Tracking Performance

To verify the online tracking performance, we consider a trajectory over which the vehicle moves towards the x^+-axis in a linear motion, turns towards the y^+-axis, keeps the linear motion, turns towards the x^--axis, and moves in a linear motion until the end. For both considered vehicle types, the velocity is constant at 11.2m/s (i.e., 25mph). Similar to the offline training data, the online radar measurements are uniformly distributed within the range of 0.2m inside the corresponding vehicle contour. As a result, the tracking performance evaluation includes the effect of spatial model mismatch as the learned spatial model only approximates the true uniform spatial model; see the notable model mismatch around the corners of Fig. 2 (b). Moreover, the number of the measurements in each scan is Poisson distributed with mean of 20,



(a) (b) (c) (d)

Fig. 3. Root mean square error (RMSE) of the UKF-PMHT algorithm with the offline learned model and the standard random matrix approach for different vehicle types: (a) position estimation for the sedan; (b) extent (length) estimation for the sedan; (c) position estimation for the truck; and (d) extent (width) estimation for the truck.

and the measurement noise variance is given as $\mathrm{diag}\left(0.1^2, 0.1^2\right)$. The sampling period is 2 seconds and there are in total 52 time steps.

For each Monte-Carlo run, the kinematic and extent states are initialized randomly as a Gaussian vector with mean $[0, 0, 11.2, 0, 0.01, 5, 2]^{\mathrm{T}}$ and covariance $\mathrm{diag}\{0.5^2, 0.5^2, 0.05, 0.1^2, 0.035^2, 0.005^2, 0.005^2\}$. For the kinematic state, the coordinated turn (CT) motion model with polar velocity [28] is used. For the extent state, i.e., the length and width, a constant model is used with a process noise with small covariance $\sigma_l^2 = \sigma_w^2 = (1e^{-5})^2$ as the physical length and width are unlikely changed over time.

Fig. 3 shows the tracking performance in terms of the root mean square error (RMSE) for the position and extent estimation for the considered vehicle types over 100 Monte-Carlo runs. The conventional RMM approach [10] is included for comparison. It is seen that, with the better learned spatial model, our UKF-PMHT algorithm can achieve lower RMSEs for both position and extent (length/width) estimations. Within the two learned models, the position errors are larger for the truck than the case of sedan, likely contributed from the larger model mismatch. The same observation also holds for the extent estimation by comparing Fig. 3 (b) and (d).

5. CONCLUSIONS

In this paper, a new surface-volume spatial model was introduced for automotive radar tracking. With an offline learned model, the UKF-PMHT algorithm has been used to deal with the nonlinear forward-warping transformation, measurement-to-ellipsis association, and kinematic/extent state update. Preliminary performance evaluation validated the proposed algorithm. Follow-up efforts include the integration of Doppler measurements [29], extension to multi-sensor fusion [22, 30] and multiple extended object tracking [31, 32], and experimental validation of real-world automotive radar measurements [33].

6. REFERENCES

[1] K. Granström, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview, and applications," *Journal of Advances in Information Fusion*, vol. 12, no. 2, 2017.

[2] P. Broßeit, M. Rapp, N. Appenrodt, and J. Dickmann, "Probabilistic rectangular-shape estimation for extended object tracking," in *2016 IEEE Intelligent Vehicles Symposium (IV)*, 2016, pp. 279–285.

[3] X. Cao, J. Lan, X. R. Li, and Y. Liu, "Extended object tracking using automotive radar," in *2018 FUSION*, 2018, pp. 1–5.

[4] M. Baum and U. D. Hanebeck, "Extended object tracking with random hypersurface models," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 50, no. 1, pp. 149–159, 2014.

[5] J.-L. Yang, P. Li, and H.-w. Ge, "Extended target shape estimation by fitting B-spline curve," *Journal of Applied Mathematics*, vol. 2014, 2014.

[6] N. Wahlström and E. Özkan, "Extended target tracking using Gaussian processes," *IEEE Trans. on Signal Processing*, vol. 63, no. 16, pp. 4165–4178, 2015.

[7] K. Thormann, M. Baum, and J. Honer, "Extended target tracking using Gaussian processes with high-resolution automotive radar," in *2018 FUSION*, 2018, pp. 1764–1770.

[8] H. Kaulbersch, J. Honer, and M. Baum, "A cartesian B-spline vehicle model for extended object tracking," in *2018 FUSION*, 2018, pp. 1–5.

[9] J. W. Koch, "Bayesian approach to extended object and cluster tracking using random matrices," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 44, no. 3, pp. 1042–1059, 2008.

[10] 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.

[11] U. Orguner, "A variational measurement update for extended target tracking with random matrices," *IEEE Trans. on Signal Processing*, vol. 60, no. 7, pp. 3827–3834, 2012.

[12] S. Yang and M. Baum, "Tracking the orientation and axes lengths of an elliptical extended object," *IEEE Trans. on Signal Processing*, vol. 67, no. 18, pp. 4720–4729, 2019.

[13] G. Yao, R. Saltus, and A. Dani, "Image moment-based extended object tracking for complex motions," *IEEE Sensors Journal*, vol. 20, no. 12, pp. 6560–6572, 2020.

[14] K. Wyffels and M. Campbell, "Precision tracking via joint detailed shape estimation of arbitrary extended objects," *IEEE Transactions on Robotics*, vol. 33, no. 2, pp. 313–332, 2016.

[15] P. Berthold, M. Michaelis, T. Luettel, D. Meissner, and H. Wuensche, "Radar reflection characteristics of vehicles for contour and feature estimation," in *2017 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*, 2017, pp. 1–6.

[16] ——, "An abstracted radar measurement model for extended object tracking," in *2018 ITSC*, 2018, pp. 3866–3872.

[17] A. Scheel and K. Dietmayer, "Tracking multiple vehicles using a variational radar model," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3721–3736, 2018.

[18] P. Broßeit, B. Duraisamy, and J. Dickmann, "The volcanormal density for radar-based extended target tracking," in *2017 IEEE ITSC*, 2017, pp. 1–6.

[19] Y. Xia, P. Wang, K. Berntorp, T. Koike-Akino, H. Mansour, M. Pajovic, P. Boufounos, and P. V. Orlik, "Extended object tracking using hierarchical truncation measurement model with automotive radar," in *2020 IEEE ICASSP*, 2020, pp. 4900–4904.

[20] Y. Xia, P. Wang, K. Berntorp, H. Mansour, P. Boufounos, and P. V. Orlik, "Extended object tracking using hierarchical truncation model with partial-view measurements," in *2020 IEEE SAM*, 2020, pp. 1–5.

[21] Y. Xia, P. Wang, K. Berntorp, P. Boufounos, P. Orlik, L. Svensson, and K. Granström, "Extended object tracking with automotive radar using learned structural measurement model," in *2020 IEEE Radar Conference*, 2020, pp. 1–6.

[22] Y. Xia, P. Wang, K. Berntorp, L. Svensson, K. Granström, H. Mansour, P. Boufounos, and P. Orlik, "Learning-based extended object tracking using hierarchical truncation measurement model with automotive rada," *IEEE Journal of Selected Topics in Signal Processing*, 2021, to appear.

[23] H. Kaulbersch, J. Honer, and M. Baum, "EM-based extended target tracking with automotive radar using learned spatial distribution models," in *2019 FUSION*, 2019, pp. 1–8.

[24] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.

[25] E. A. Wan and R. Van Der Merwe, "The unscented Kalman filter for nonlinear estimation," in *IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium*, 2000, pp. 153–158.

[26] C. de Boor, *A practical guide to splines*. Springer-Verlag New York, 1978, vol. 27.

[27] M. Wieneke and W. Koch, "A PMHT approach for extended objects and object groups," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 3, pp. 2349–2370, 2012.

[28] X. Rong Li and V. P. Jilkov, "Survey of maneuvering target tracking. part I. dynamic models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1333–1364, 2003.

[29] M. Schuster, J. Reuter, and G. Wanielik, "Probabilistic data association for tracking extended targets under clutter using random matrices," in *2015 Fusion*, 2015, pp. 961–968.

[30] G. Vivone, K. Granström, P. Braca, and P. Willett, "Multiple sensor measurement updates for the extended target tracking random matrix model," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2544–2558, 2017.

[31] K. Granström and U. Orguner, "A PHD filter for tracking multiple extended targets using random matrices," *IEEE Trans. on Signal Processing*, vol. 60, no. 11, pp. 5657–5671, 2012.

[32] K. Granström, M. Fatemi, and L. Svensson, "Poisson multi-Bernoulli mixture conjugate prior for multiple extended target filtering," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 56, pp. 208–225, 2019.

[33] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuScenes: A multimodal dataset for autonomous driving," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11 621–11 631.