Joint Tracking and Classification of Vehicles with the PHD Filter and Gaussian Processes

Jiaye Yang, Yuhuan Xiong, Xi Cao, Cong Peng, Wei Yi
School of Information and Communication Engineering
University of Electronic Science and Technology of China, Chengdu, China
Email: {yangjy991105, xyh202322, caoxi0419, panycg_cn}@163.com, kussoyi@gmail.com

***Abstract*—Joint tracking and classification (JTC) of vehicles is a crucial yet challenging task in intelligent transport and automotive systems. The advent of high-resolution modern sensors necessitates treating vehicles as extended targets. Current extended target tracking (ETT) algorithms provide shape estimations for vehicles, making shape size the most intuitive and accessible feature for classification. This paper contributes two key elements to achieve the JTC of vehicles. For one thing, we introduce the rectangular constraints and customize distinguishable measurement models using modified Gaussian processes (GP). For another thing, based on the customized GP models, we strengthen the role of class in the conditional extended target probability hypothesis density (ET-PHD) filter. Subsequently, we propose a class-enhanced JTC-ET-PHD filter and its Gaussian mixture implementation, enabling simultaneous kinematic, shape, and class estimation of vehicles. Finally, numerical results validate the proposed shape estimation and JTC method, affirming their effectiveness in addressing JTC challenges.**

***Index Terms*—JTC, vehicle tracking and classification, extended target, Gaussian processes, PHD filter.**

I. Introduction

In recent years, the development of automated vehicles and intelligent transport systems has surged with the integration of modern sensors. The tracking and classification of traffic targets are the issues of great concern for these systems [1, 2]. Meanwhile, the modern sensors, like automotive radar and laser, can produce plenty of measurements just from one target due to their high resolution [3]. The target cannot be viewed as a point source, but an object with a certain shape instead. That is also known as extended target. Beyond capturing the kinematic information and trajectory of vehicles, current manufacturers also want to obtain their precise shapes and categories. So, the joint tracking and classification (JTC) has to be upgraded to fulfill the application requirements.

The original JTC theory [4] was derived within a Bayesian framework in the early years. Nevertheless, subject to the limited performance of sensor at that time, the targets could only be perceived as point sources. There were limited features that can be used for classification. In [5] and [6], the high resolution range profile (HRRP) and radar cross section (RCS) were first introduced as the primary classification bases for JTC. Notably, those properties are highly sensitive to the pose of target. The changes in pose may cause intense fluctuations in the target echos. Since the tracking algorithms can estimate the kinematic information of targets, some researches explore the use of flight envelope, velocity, motion model and so on for target class determination [7, 8]. However, if the targets have similar RCS and kinematic states, those methods will fail to achieve effective JTC. In fact, that is not an uncommon occurrence, especially in traffic scenes, where the vehicles cannot be distinguished by the attributes mentioned above.

Fortunately, the modern automotive and roadside sensors have higher resolution. The vehicles are regarded as extended targets. The numerous measurements they generate can even support filters to estimate their shapes. So, obtaining the precise shape sizes of vehicles allows for intuitive recognition, for instance, distinguishing between a car and a truck. At present, there are a wealth of researches on extended target tracking (ETT) and shape estimation, in which a reasonable shape model is very important. Prior shape information, like ellipse, has been considered in [9–11]. The random matrix model (RMM) [9] for elliptical shape is very popular, which has been realized in many advanced filters [12–14]. In addition to the fixed shapes, the star-convex model defines the radius connecting the center and the convex contour through a radial function. Both the random hypersurface model (RHM) [15] and Gaussian processes (GP) regression [16] are based on the star-convex shape. The RHM decomposes the unknown radial function with a set of Fourier coefficients, while the GP constructs a radial function and a covariance function in the spatial domain. The GP method is more intuitive and retains the uncertainty of extent at each scan, which is congenial to most ETT filters. The GP regression has been successfully implemented in vehicle tracking [17–19]. To improve the efficiency, our previous work integrates prior rectangular information into GP regression [20].

Hence, substantial improvement can be anticipated by combining the JTC method with vehicle shape estimation. Nonetheless, although existing related algorithms are not uncommon, they are not well suited for vehicle tracking and classification. In [21] and [22], the JTC of extended targets is realized with RMM. However, the elliptical extension or shape approximation of multiple ellipses may not be very suitable for vehicles. For star-convex shapes, the JTC theory and RHM have been integrated into the probability

This work was supported in part by the National Natural Science Foundation of China under Grant 62231008 and 62301127, in part by the China Postdoctoral Science Foundation under Grant BX20220057, 2023M730509 and GZB20230112, the "Tianfu Qingcheng" Plan of Sichuan Province under Grant1332 and 1395.

hypothesis density (PHD) [23] and cardinality balanced multi-target multi-Bernoulli (CBMeMBer) filters [24]. In [25], a stepwise ETT and shape classification are accomplished with the GP regression, which, strictly speaking, is not a JTC algorithm. Although the classification methods based on star-convex models [23–25] can classify targets with similar sizes but different shapes, they are not applicable in traffic scenes because vehicles tend to have similar rectangular shapes but varying sizes. Besides, most of the existing JTC methods for extended targets [21–24] require the inclusion of prior class-dependent information, also called as pseudo-measurements, to derive differentiated class probabilities.

In this paper, we aim to explore a new JTC method for rectangular vehicles without additional pseudo-measurements. Our approach enables the simultaneous extraction of kinematic, shape, and class information as long as feeding the proposed filter with the raw measurements and necessary parameters of customized multiple models. The contributions of this work can be summarized as follows:

- For different classes of vehicles, we customize distinguishable GP models with soft rectangular constraints. That modeling approach stems from our preliminary work [20]. We develop it into a more refined version for different shape sizes, which can be beneficial to classification even though the vehicles have similar shapes.
- A class-enhanced JTC-ET-PHD filter is proposed without the necessity for additional pseudo-measurements. By embedding each customized GP model into the corresponding sub-filter, we strengthen the constraint of class in the filtering. Detailed filtering formulas and Gaussian mixture implementation will be derived in this paper.

II. Extended Target Modeling

In our effort to classify rectangular traffic targets just in accordance with their shape sizes, we first need estimate the kinematic and extent states jointly. This section will make a brief review of the GP model [16] for star-convex targets and our modification on introducing some soft rectangular constraints [20].

A. Star-convex Extent Model

The star-convex shapes are sculptured with a number of basic angles $\mathbf{u}^{\mathrm{f}} \triangleq [u_1^{\mathrm{f}}, ..., u_{N^{\mathrm{f}}}^{\mathrm{f}}]^\top$ and the corresponding radii $f(\mathbf{u}^{\mathrm{f}}) \triangleq [f(u_1^{\mathrm{f}}), ..., f(u_{N^{\mathrm{f}}}^{\mathrm{f}})]^\top$. Each radius connects the center with the convex contour. It should be noted that the radial function $f(\mathbf{u}^{\mathrm{f}})$ with a period of 2π can be arbitrary depending on the real shape. For better shape estimation, of course, the basic angles and radial function can be elaborately designed if some prior knowledge is available. For example, vehicles are usually described as rectangles in traffic tracking scenes.

Fig. 1(a) is a general demonstration for the following comprehensible mathematical explanation. Firstly, the target position, orientation and radii are defined by $\mathbf{x}_k^c$, ψ_k and $\mathbf{x}_k^{\mathrm{f}} \triangleq f(\mathbf{u}^{\mathrm{f}})$ at the timescale k, respectively. We set the target state as $\mathbf{x}_k \triangleq [(\bar{\mathbf{x}}_k)^\top, (\mathbf{x}_k^{\mathrm{f}})^\top]^\top$, where $\bar{\mathbf{x}}_k$ denotes the kinematic state, defined by $\bar{\mathbf{x}}_k \triangleq [(\mathbf{x}_k^c)^\top, \psi_k, (\mathbf{x}_k^*)^\top]^\top$. The

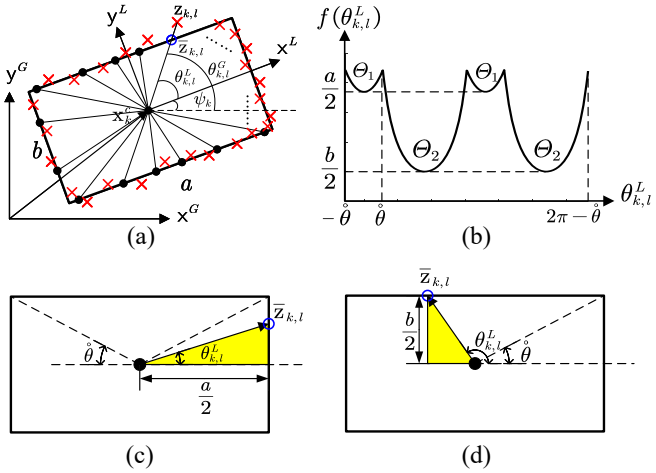

Fig. 1. A star-convex target is composed of a center $\mathbf{x}_k^c$ and many radii functioned by local angles $\theta_{k,l}^L$. In Fig. 1(a), the red crosses represent the measurements and the blue circle is the mean of one measurement among them. The notations x^G, y^G and x^L, y^L denote the global and local polar coordinates, respectively. The Fig. 1(b) is the curve of rectangular radial function $f(\theta_{k,l}^L)$. The Fig. 1(c) and Fig. 1(d) show two cases of the radial function in different angular scopes. In order to get $f(\theta_{k,l}^L)$, we can perform some geometry and trigonometry in these yellow triangles, where $f(\theta_{k,l}^L)$ is just the length of the hypotenuse.

additional state $\mathbf{x}_k^*$ is chosen as velocity and yaw rate in this paper. Meanwhile, an extended target may generate a cluster of measurements $\mathbf{z}_k \triangleq [\mathbf{z}_{k,1}^\top, ..., \mathbf{z}_{k,n_k}^\top]^\top$ around its extent. In the star-convex model, each measurement $\mathbf{z}_{k,l}$ generates an angle $\theta_{k,l}^G$ in global polar coordinates and a corresponding angle $\theta_{k,l}^L$ in local polar coordinates. The angles are generated by

$$\theta_{k,l}^G(\mathbf{x}_k^c) = \angle(\mathbf{z}_{k,l} - \mathbf{x}_k^c),\ \theta_{k,l}^L(\mathbf{x}_k^c, \psi_k) = \theta_{k,l}^G(\mathbf{x}_k^c) - \psi_k. \quad (1)$$

Next, with the help of the radial function, the measurement equation can be described as

$$\begin{aligned} \mathbf{z}_{k,l} &= \bar{\mathbf{z}}_{k,l} + \bar{\mathbf{e}}_{k,l} \\ &= \mathbf{x}_k^c + \mathbf{p}_{k,l}\left(\theta_{k,l}^G\right) \cdot f\left(\theta_{k,l}^L\right) + \bar{\mathbf{e}}_{k,l}, \end{aligned} \quad (2)$$

where $\bar{\mathbf{z}}_{k,l}$ denotes the mean value; $\bar{\mathbf{e}}_{k,l}$ is the measurement noise following $\bar{\mathbf{e}}_{k,l} \sim \mathcal{N}\left(0, \bar{R}_{k,l}\right)$ and $\mathbf{p}_{k,l}$ is the orientation vector given by $\mathbf{p}_{k,l}(\theta_{k,l}^G) = [\cos\theta_{k,l}^G, \sin\theta_{k,l}^G]^\top$. Then, we can just portray the arbitrary shape of target with a number of basic angles and radii.

B. Recursive Gaussian Processes Regression

In practice, characterized by the radial function $f(\theta_{k,l}^L)$ in (2), the shape of target tends to be uncertain. The recursive GP regression is primarily suitable for training, learning and modeling such an unknown function. In [16], it has been applied in ETT and shape estimation. Further in our previous work [20], the radial function is modeled by a non-zero mean GP for rectangular vehicles, given by

$$f(\theta) \sim \mathcal{GP}\left(\mu(\theta), k(\theta, \theta')\right). \quad (3)$$

The mean function $\mu(\theta)$ is calculated by

$$\mu(\theta)=\begin{cases}\dfrac{a}{2|\cos\theta|}, & \text{if} \quad \theta\in\Theta_1\\ \dfrac{b}{2|\sin\theta|}, & \text{if} \quad \theta\in\Theta_2\end{cases}, \tag{4}$$

where a, b denote the length and width of the rectangle; $|\cdot|$ means taking the absolute value; the scope of conditions Θ_1 and Θ_2 can be described as

$$\begin{cases}\Theta_1=\{\theta \mid \mathrm{n}\pi-\mathring{\theta}\leqslant\theta\leqslant\mathrm{n}\pi+\mathring{\theta}\}\\ \Theta_2=\{\theta \mid \mathrm{n}\pi+\mathring{\theta}\leqslant\theta\leqslant(\mathrm{n}+1)\pi-\mathring{\theta}\}\end{cases}; \tag{5}$$

$\mathring{\theta}$ is defined by $\mathring{\theta}=arctan(b/a)$; the number n is an integer characterizing the periodicity. The graphical interpretation is shown in the Fig. 1. The radial function will be learned during the filtering. In (3), the variance function $k(\theta,\theta')$ is a periodic kernel function with a form of

$$k(\theta,\theta')=\sigma_f^2 e^{-\frac{2\sin^2\left(\frac{|\theta-\theta'|}{2}\right)}{\ell^2}}+\sigma_r^2, \tag{6}$$

where σ_f^2 is the prior variance, ℓ denotes the length scale of the radial function and σ_r^2 is the variance of the radius. In [16], the likelihood function is built as a zero mean GP, which can be easily generalized to a non-zero mean version, given by

$$\phi_{\mathbf{z}_{k,l}}(\mathbf{x}_k)=\mathcal{N}(\mathbf{z}_{k,l};h_{k,l}(\mathbf{x}_k),R_{k,l}), \tag{7a}$$

$$h_{k,l}(\mathbf{x}_k)=\mathbf{x}_k^c+\mathbf{p}_{k,l}\left(\mu(\theta_{k,l}^L)+H^{\mathrm{f}}(\theta_{k,l}^L)\cdot(\mathbf{x}_k^{\mathrm{f}}-\mu(\mathbf{u}^{\mathrm{f}}))\right), \tag{7b}$$

$$R_{k,l}=\mathbf{p}_{k,l}R^{\mathrm{f}}(\theta_{k,l}^L)\mathbf{p}_{k,l}^\top+\bar{R}_{k,l}, \tag{7c}$$

where $H^{\mathrm{f}}(\cdot)$ and $R^{\mathrm{f}}(\cdot)$ are measurement and covariance functions, defined by

$$H^{\mathrm{f}}(\theta_{k,l}^L)=K(\theta_{k,l}^L,\mathbf{u}^{\mathrm{f}})[K(\mathbf{u}^{\mathrm{f}},\mathbf{u}^{\mathrm{f}})]^{-1}, \tag{8}$$

$$R^{\mathrm{f}}(\theta_{k,l}^L)=k(\theta_{k,l}^L,\theta_{k,l}^L)-K(\theta_{k,l}^L,\mathbf{u}^{\mathrm{f}})[K(\mathbf{u}^{\mathrm{f}},\mathbf{u}^{\mathrm{f}})]^{-1}K(\mathbf{u}^{\mathrm{f}},\theta_{k,l}^L). \tag{9}$$

The function $K(\cdot)$ is the matrix form of $k(\cdot)$, given by

$$K\left(\mathbf{u},\mathbf{u}^{\mathrm{f}}\right)=\begin{bmatrix}k\left(u_1,u_1^{\mathrm{f}}\right) & \dots & k\left(u_1,u_{N^{\mathrm{f}}}^{\mathrm{f}}\right)\\ \vdots & \ddots & \vdots\\ k\left(u_N,u_1^{\mathrm{f}}\right) & \dots & k\left(u_N,u_{N^{\mathrm{f}}}^{\mathrm{f}}\right)\end{bmatrix}. \tag{10}$$

At last, in the light of the random finite set (RFS) filter, the RFSs of target states and measurements should be defined by

$$\mathbf{X}_k=\left\{\mathbf{x}_k^{(1)},\mathbf{x}_k^{(2)},\dots\right\},\quad \mathbf{Z}_k=\left\{\mathbf{z}_k^{(1)},\mathbf{z}_k^{(2)},\dots\right\}. \tag{11}$$

III. Joint Tracking and Classification of Vehicle Extended Targets

In this section, we will construct multiple rectangular GP models, each of which needs to accommodate a degree of uncertainty but be distinguishable from others. Then, a class-enhanced JTC-ET-PHD (CE-JTC-ET-PHD) filter will be derived based on the multiple tailor-made GP models.

A. Multiple Rectangular GP Models for Classification

Given our emphasis on using shape size as the classification indicator, it is imperative to customize the extent models of vehicles to suit different classes. Firstly, we define the class set with N_c classes as $\boldsymbol{C}=\left\{c^{(i)}\right\}_{i=1}^{N_c}$. The target class c_k takes value in the predefined class set $\boldsymbol{C}$. Representing the characteristics of various traffic targets, the basic angles may vary among multiple classes. To show the difference, the class index i is introduced into the definition of basic angles, given by $\mathbf{u}^{\mathrm{f}(i)}\triangleq[u_1^{\mathrm{f}(i)},...,u_{N^{\mathrm{f}(i)}}^{\mathrm{f}(i)}]^\top$.

As the extent model has been built with rectangular information and GP in Section II, different lengths and widths with preconfigured uncertainties should be considered for each class $c^{(i)}$. Therefore, we redefine $\mu(\theta)$ in (4) with class index i as

$$\mu^{(i)}(\theta)=\begin{cases}\dfrac{a^{(i)}}{2|\cos\theta|}, & \text{if} \quad \theta\in\Theta_1^{(i)}\\ \dfrac{b^{(i)}}{2|\sin\theta|}, & \text{if} \quad \theta\in\Theta_2^{(i)}\end{cases}, \tag{12}$$

where $a^{(i)}$ and $b^{(i)}$ denote the length and width in the i-th class; the scope of conditions $\Theta_1^{(i)}$ and $\Theta_2^{(i)}$ can be obtained by replacing $\mathring{\theta}$ with $\mathring{\theta}^{(i)}$ in (5); $\mathring{\theta}^{(i)}$ is given by $\mathring{\theta}^{(i)}=arctan(b^{(i)}/a^{(i)})$. We also customize the uncertainties for different classes. Let $\sigma_a^{2(i)}$ and $\sigma_b^{2(i)}$ denote the variance of the length and width in the i-th class. Thus, the variance of radius in a specific class can be redefined as

$$\sigma_r^{2(i)}=\begin{cases}\dfrac{\sigma_a^{2(i)}}{4|\cos\theta||\cos\theta'|}, & \text{if} \quad \theta\in\Theta_1^{(i)},\theta'\in\Theta_1^{(i)}\\ \dfrac{\sigma_a^{(i)}\sigma_b^{(i)}}{4|\cos\theta||\sin\theta'|}, & \text{if} \quad \theta\in\Theta_1^{(i)},\theta'\in\Theta_2^{(i)}\\ \dfrac{\sigma_b^{(i)}\sigma_a^{(i)}}{4|\sin\theta||\cos\theta'|}, & \text{if} \quad \theta\in\Theta_2^{(i)},\theta'\in\Theta_1^{(i)}\\ \dfrac{\sigma_b^{2(i)}}{4|\sin\theta||\sin\theta'|}, & \text{if} \quad \theta\in\Theta_2^{(i)},\theta'\in\Theta_2^{(i)}\end{cases}. \tag{13}$$

Substitute (13) into (6), and we can get the specific kernel function in the i-th class, defined by

$$k^{(i)}(\theta,\theta')=\sigma_f^2 e^{-\frac{2\sin^2\left(\frac{|\theta-\theta'|}{2}\right)}{\ell^2}}+\sigma_r^{2(i)}. \tag{14}$$

Correspondingly, the function $K^{(i)}(\cdot)$ can be calculated by (10). With the redefined mean and covariance functions in (12) and (14), the customized radial function for the i-th class can be represented by a conditional function on the class $c^{(i)}$,

$$f(\theta \mid c^{(i)})\sim\mathcal{GP}\left(\mu^{(i)}(\theta),k^{(i)}(\theta,\theta')\right), \tag{15}$$

where it should be noted that $\mu^{(i)}(\theta)$ and $k^{(i)}(\theta,\theta')$ are equivalent to $\mu(\theta|c^{(i)})$ and $k(\theta,\theta'|c^{(i)})$ although their symbolic expressions are different. Therefore, the likelihood function is also conditioned on the class $c^{(i)}$, given by

$$\phi_{\mathbf{z}_{k,l}}(\mathbf{x}_k|c^{(i)})=\mathcal{N}\left(\mathbf{z}_{k,l};h_{k,l}(\mathbf{x}_k|c^{(i)}),R_{k,l}^{(i)}\right), \tag{16a}$$

$$h_{k,l}(\mathbf{x}_k|c^{(i)})=\mathbf{x}_k^c+\mathbf{p}_{k,l}\cdot\mu^{(i)}(\theta_{k,l}^L)+\mathbf{p}_{k,l}\cdot H^{\mathrm{f}(i)}(\theta_{k,l}^L)\cdot(\mathbf{x}_k^{\mathrm{f}}-\mu^{(i)}(\mathbf{u}^{\mathrm{f}(i)})), \tag{16b}$$

$$R_{k,l}^{(i)}=\mathbf{p}_{k,l}R^{\mathrm{f}(i)}(\theta_{k,l}^L)\mathbf{p}_{k,l}^\top+\bar{R}_{k,l}, \tag{16c}$$

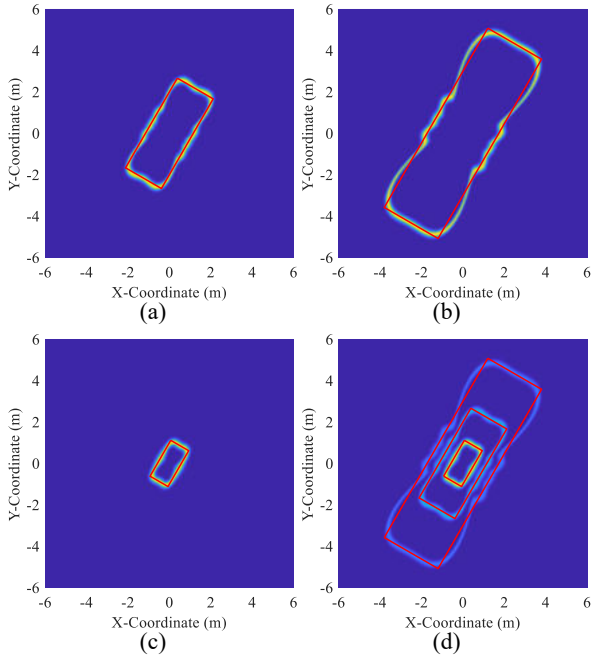

Fig. 2. Some examples of the GP likelihoods in different classes. The red boxes represent the real shapes of vehicles. The Fig. 2(a-c) are car, truck and bicycle model, respectively. In Fig. 2(d), they are aggregated in a single graph.

where $H^{\mathrm{f}(i)}(\cdot)$ and $R^{\mathrm{f}(i)}(\cdot)$ can be obtained by replacing $\mathbf{u}^{\mathrm{f}}$ with $\mathbf{u}^{\mathrm{f}(i)}$ in (8) and (9). In Fig. 2, we give some examples of the GP likelihoods in different classes. The shape size settings are the same as those in Table I. In Fig. 2(d), we can find that they do not interfere with each other and are clearly recognizable. That is also essentially why we can conduct the classification successfully.

B. Class-Enhanced JTC-ET-PHD Filter

We expect to strengthen the role of class in constraining filtering given the distinguishable GP models established in Section III.A. In this section, we will design a CE-JTC-ET-PHD filter for different classes of vehicles.

1) CE-JTC-ET-PHD prediction: The prediction step of our proposed filter is similar to the normal JTC-PHD filter in [26], which obeys the Chapman-Kolmogorov equation, given by

$$D_{k|k-1}(\mathbf{x},c)=\iint B_k(\mathbf{x},c')f_{k|k-1}(c|c')+p_{\mathrm{S}}(\mathbf{x},c') \times f_{k|k-1}(\mathbf{x},c|\mathbf{x}',c')D_{k-1}(\mathbf{x}',c')d\mathbf{x}'dc', \quad (17)$$

where $B_k(\cdot)$ is the birth density and $p_{\mathrm{S}}(\cdot)$ is the survival probability. We assume that the class is independent of the augmented state in the transition density. So, the transition density and the former PHD can be factorized as

$$f_{k|k-1}(\mathbf{x},c|\mathbf{x}',c')=f_{k|k-1}(\mathbf{x}|\mathbf{x}',c)f_{k|k-1}(c|c'), \quad (18)$$

$$D_{k-1}(\mathbf{x}',c')=D_{k-1}(\mathbf{x}'|c')p_{k-1}(c'), \quad (19)$$

where $f_{k|k-1}(\mathbf{x}|\mathbf{x}',c)$ means the transition density of target state under the condition of the known class c; $f_{k|k-1}(c|c')$ is the class transition probability, which takes value in the transition probability matrix (TPM) defined by $T_{c|c'}$,

$$f_{k|k-1}(c|c')=[T_{c|c'}]_{(i,j)}=Pr[c_k=c^{(i)}|c_{k-1}=c^{(j)}]. \quad (20)$$

The notation $[\cdot]_{(i,j)}$denotes the element at the i-th row and j-th column of the matrix. Substitute (18-20) into (17), and we can get the prediction equation. The predicted class probability $p_{k|k-1}(c)$, representing $Pr[c_{k|k-1}=c^{(i)}]$, is calculated by

$$p_{k|k-1}(c)=[T_{c|c'}\pi_{k-1}(c')]_{(i)}/\sum_{j=1}^{N_c}[T_{c|c'}\pi_{k-1}(c')]_{(j)}, \quad (21)$$

where $\pi_{k-1}(c')$ denotes the probability distribution of all the classes at the time step $k-1$, defined by $\pi_{k-1}(c')=[p_{k-1}(c^{(1)}),p_{k-1}(c^{(2)}),\ldots,p_{k-1}(c^{(N_c)})]^{\top}$; the notation $[\cdot]_{(i)}$ denotes the i-th element of the vector.

2) CE-JTC-ET-PHD update: Firstly, the posterior PHD of joint target state and class can be factorized as

$$D_{k|k}(\mathbf{x},c|\mathbf{Z})=p_{k|k}(c|\mathbf{Z})D_{k|k}(\mathbf{x}|c,\mathbf{Z}). \quad (22)$$

Since we emphasize the constraining role of class in the filtering, the likelihood function should be conditioned on the class during the update step,

$$D_{k|k}(\mathbf{x}|c,\mathbf{Z})=L_{\mathbf{Z}_k}(\mathbf{x}|c)D_{k|k-1}(\mathbf{x}|c). \quad (23)$$

Evolved from the normal ET-PHD filter in [27, 28], the class-conditioned likelihood function can be given by

$$L_{\mathbf{Z}_k}(\mathbf{x}|c)=1-\left(1-e^{-\gamma(\mathbf{x}|c)}\right)p_{\mathrm{D}}(\mathbf{x}|c)+e^{-\gamma(\mathbf{x}|c)}p_{\mathrm{D}}(\mathbf{x}|c) \times\sum_{\mathcal{P}\angle\mathbf{Z}_k}\omega_{\mathcal{P}}(c)\sum_{W\in\mathcal{P}}\frac{\gamma(\mathbf{x}|c)^{|W|}}{d_W(c)}\prod_{\mathbf{z}_{k,l}\in W}\frac{\phi_{\mathbf{z}_{k,l}}(\mathbf{x}|c)}{\lambda_k\mathcal{C}(\mathbf{z}_{k,l})}, \quad (24)$$

where $\gamma(\cdot)$ is the measurement rate; $p_{\mathrm{D}}(\cdot)$ is the detection probability; $|W|$ denotes the number of measurements in the cluster W; the single measurement likelihood function $\phi_{\mathbf{z}_{k,l}}(\mathbf{x}|c)$ has been given in (16), using customized GP models; λ_k is the Poisson clutter intensity; $\mathcal{C}(\mathbf{z}_{k,l})$ is the spatial distribution of clutters. Both $\omega_{\mathcal{P}}(c)$ and d_W are normalization coefficients. According to the Bayesian formula, the posterior class probability can be calculated by

$$p_{k|k}(c|\mathbf{Z})=\frac{L_{\mathbf{Z}_k}(c)p_{k|k-1}(c)}{\sum_{i=1}^{N_c}L_{\mathbf{Z}_k}(c^{(i)})p_{k|k-1}(c^{(i)})}. \quad (25)$$

Thus, we obtain the update equations for the extended target conditional PHD in (23) and the class probability in (25), respectively. Then, the marginal posterior PHD can be derived by a mix operation involving all the classes,

$$D_{k|k}(\mathbf{x})=\sum_{i=1}^{N_c}D_{k|k}(\mathbf{x}|c^{(i)},\mathbf{Z})p_{k|k}(c^{(i)}|\mathbf{Z}). \quad (26)$$

Meanwhile, we can judge the class of the target using the MAP principle,

$$c_k=arg\max_{c^{(i)}\in\boldsymbol{C}}\left[p_{k|k}(c^{(i)}|\mathbf{Z})\right]. \quad (27)$$

C. Gaussian Mixture Implementation

For the implementation of proposed filter, Gaussian mixture approximation proves to be a simple but effective method in [28, 29]. According to the earlier derivation, the filtering process of CE-JTC-ET-PHD is segmented into two parts: the conditional PHD and class probability. In this context, the class-dependent Gaussian mixtures are employed to approximate the conditional PHD.

If we define the mean and covariance of the Gaussian component as $\Theta \triangleq \{m, P\}$, the prior conditional PHD can be represented by

$$D_{k|k-1}(\mathbf{x}|c) = \sum_{j=1}^{J_{\mathrm{B},k}} w_{\mathrm{B},k}^{(j)}(c)\mathcal{N}\left(\mathbf{x}|c;\Theta_{\mathrm{B},k}^{(j)}(c)\right) + \sum_{j=1}^{J_{k-1}} w_{\mathrm{S},k|k-1}^{(j)}(c)\mathcal{N}\left(\mathbf{x}|c;\Theta_{\mathrm{S},k|k-1}^{(j)}(c)\right), \quad (28)$$

where J and w denote the number and weight of Gaussian components, respectively; the superscript (j) denotes the j-th Gaussian component; the identifier "B" and "S" denote the birth and survival components. Similarly, the posterior conditional PHD can also have a form of Gaussian mixtures,

$$D_{k|k}(\mathbf{x}|c,\mathbf{Z}) = \sum_{j=1}^{J_{k|k-1}} w_{\mathrm{ND},k|k}^{(j)}(c)\mathcal{N}\Big(\mathbf{x}|c;\Theta_{\mathrm{ND},k|k}^{(j)}(c)\Big) + \sum_{\mathcal{P}\angle\mathbf{Z}_k}\sum_{W\in\mathcal{P}}\sum_{j=1}^{J_{k|k-1}} w_{\mathrm{D},k|k}^{W,(j)}(c)\mathcal{N}\Big(\mathbf{x}|c;\Theta_{\mathrm{D},k|k}^{W,(j)}(c)\Big), \quad (29)$$

where the identifiers "ND" and "D" denote the undetected and detected components. Then, we make the following approximations,

$$p_{\mathrm{S}}(\mathbf{x}|c) \approx p_{\mathrm{S}}^{(j|c)},\ e^{-\gamma(\mathbf{x}|c)} \approx e^{-\gamma^{(j|c)}},\ p_{\mathrm{D}}(\mathbf{x}|c) \approx p_{\mathrm{D}}^{(j|c)}, \quad (30)$$

where we use the abbreviations $p_{\mathrm{S}}^{(j|c)}$, $\gamma^{(j|c)}$ and $p_{\mathrm{D}}^{(j|c)}$ for $p_{\mathrm{S}}(m^{(j)}|c)$, $\gamma(m^{(j)}|c)$ and $p_{\mathrm{D}}(m^{(j)}|c)$, respectively.

1) GM implementation for prediction: In (28), the survival terms can be obtained by

$$w_{\mathrm{S},k|k-1}^{(j)}(c) = p_{\mathrm{S}}^{(j|c)} w_{k-1}^{(j)}(c), \quad (31)$$

$$\Theta_{\mathrm{S},k|k-1}^{(j)}(c) = UT\left(f_{k|k-1}(\mathbf{x}|\mathbf{x}',c);\Theta_{k-1}^{(j)}(c)\right), \quad (32)$$

where $UT(\cdot)$ denotes unscented transformation (UT) [29–31], handling the possible nonlinear state transition. Meanwhile, the prediction equation of class probability has been given in (21), which contains a class transition and a mix operation.

2) GM implementation for update: By substituting the predicted GMs (28) into (23), we can derive the calculation formulas of the undetected and detected terms. In (29), the weights, means and covariances of the undetected components can be calculated by

$$w_{\mathrm{ND},k|k}^{(j)}(c) = \left(1 - \left(1 - e^{-\gamma^{(j|c)}}\right) p_{\mathrm{D}}^{(j|c)}\right) w_{k|k-1}^{(j)}(c), \quad (33)$$

$$\Theta_{\mathrm{ND},k|k}^{(j)}(c) = \Theta_{k|k-1}^{(j)}(c). \quad (34)$$

In (29), the weights of detected components are calculated by

$$w_{\mathrm{D},k|k}^{W,(j)}(c) = \omega_{\mathcal{P}}(c)\frac{\Gamma^{(j|c)}\Phi_k^{W,(j)}(c)}{d_W(c)\prod_{\mathbf{z}_{k,l}\in W}\lambda_k\mathcal{C}(\mathbf{z}_{k,l})} w_{k|k-1}^{(j)}(c), \quad (35a)$$

$$\Gamma^{(j|c)} = e^{-\gamma^{(j|c)}}(\gamma^{(j|c)})^{|W|} p_{\mathrm{D}}^{(j|c)}, \quad (35b)$$

$$\Phi_k^{W,(j)}(c) = \prod_{\mathbf{z}_{k,l}\in W} \phi_{\mathbf{z}_{k,l}}(\mathbf{x}|c) = \mathcal{N}\left(\mathbf{z}_k^W; \bar{\mathbf{z}}_{k|k-1}^{W,(j)}(c), S_{k|k}^{W,(j)}(c)\right), \quad (35c)$$

$$\omega_{\mathcal{P}}(c) = \frac{\prod_{W\in\mathcal{P}} d_W(c)}{\sum_{\mathcal{P}'\angle\mathbf{Z}_k}\prod_{W'\in\mathcal{P}'} d_{W'}(c)}, \quad (35d)$$

$$d_W(c) = \delta_{|W|,1} + \sum_{j=1}^{J_{k|k-1}} \frac{\Gamma^{(j|c)}\Phi_k^{W,(j)}(c)}{\prod_{\mathbf{z}_{k,l}\in W}\lambda_k\mathcal{C}(\mathrm{z}_{k,l})} w_{k|k-1}^{(j)}(c), \quad (35e)$$

where $\delta_{|W|,1}$ is the Kronecker delta; $\mathbf{z}_k^W$ is a cluster of measurements representing $[\mathbf{z}_{k,1}^\top, ..., \mathbf{z}_{k,|W|}^\top]^\top$; the likelihood (35c) for $\mathbf{z}_k^W$ can be derived by multiplying the single measurement likelihoods in (16). As it is a nonlinear likelihood function, the mean and covariance can be obtained through a UT process,

$$\left\{\bar{\mathbf{z}}_{k|k-1}^{W,(j)}(c), S_{k|k}^{W,(j)}(c), \Theta_{\mathrm{D},k|k}^{W,(j)}(c)\right\} = UT\left(h_k^W(\mathbf{x}|c); \Theta_{k|k-1}^{(j)}(c)\right), \quad (36)$$

where $h_k^W(\mathbf{x}|c)$ is the vertical augmented form of $h_{k,l}(\mathbf{x}|c)$. Details of UT can be found in [29–31] and our previous work in [20, Algorithm 1]. On the other hand, the posterior class probability $p_{k|k}(c|\mathbf{Z})$ can also be calculated by

$$p_{k|k}(c|\mathbf{Z}) = \left\{p_{\mathrm{ND},k|k}^{(j)}(c|\mathbf{Z})\right\}_{j=1}^{J_{k|k-1}} \bigcup_{\mathcal{P}\angle\mathbf{Z}_k, W\in\mathcal{P}} \cdots \bigcup \left\{\left\{p_{\mathrm{D},k|k}^{W,(j)}(c|\mathbf{Z})\right\}_{j=1}^{J_{k|k-1}}\right\}, \quad (37)$$

$$p_{\mathrm{ND},k|k}^{(j)}(c|\mathbf{Z}) = \left(1 - \left(1 - e^{-\gamma^{(j|c)}}\right) p_{\mathrm{D}}^{(j|c)}\right) p_{k|k-1}^{(j)}(c) \Big/ \varpi_{\mathrm{ND}}, \quad (38)$$

$$p_{\mathrm{D},k|k}^{W,(j)}(c|\mathbf{Z}) = d_W(c) p_{k|k-1}^{(j)}(c) \Big/ \varpi_{\mathrm{D}}, \quad (39)$$

where ϖ_{ND} and ϖ_{D} are normalization factors, calculated by

$$\varpi_{\mathrm{ND}} = \sum_{i=1}^{N_c}\left(1 - \left(1 - e^{-\gamma^{(j|c^{(i)})}}\right) p_{\mathrm{D}}^{(j|c^{(i)})}\right) p_{k|k-1}^{(j)}(c^{(i)}), \quad (40)$$

$$\varpi_{\mathrm{D}} = \sum_{i=1}^{N_c} d_W(c^{(i)}) p_{k|k-1}^{(j)}(c^{(i)}). \quad (41)$$

Then, if we want to get the GMs of marginal posterior PHD, they can be acquired by a mix operation,

$$\left\{w_{k|k}^{(j)}, \Theta_{k|k}^{(j)}\right\} = \sum_{i=1}^{N_c} p_{k|k}^{(j)}(c^{(i)}|\mathbf{Z})\left\{w_{k|k}^{(j)}(c^{(i)}), \Theta_{k|k}^{(j)}(c^{(i)})\right\}. \quad (42)$$

IV. Numerical Results

In this paper, the scenario is set up with four vehicles moving in a constant velocity (CV) motion at an intersection. We consider 3 classes of common vehicles, as shown in Table

TABLE I
THE CLASS PARAMETERS OF THE SIMULATION.

class $c^{(i)}$	length $a^{(i)}$	width $b^{(i)}$	standard deviation $\sigma_a^{(i)}$	$\sigma_b^{(i)}$	measurement rate $\gamma(\cdot\|c^{(i)})$
1	5m	2m	0.6m	0.2m	10
2	10m	3m	2m	0.2m	10
3	2m	1m	0.2m	0.1m	5

TABLE II
THE INITIAL STATES AND TIME CHART OF TARGETS.

target	initial state	birth	death
T1	[-60m;57.5m;0°;5m/s;0m/s;0rad/s;$\mu(\theta\|c^{(3)})$]	0.2s	10s
T2	[60m;62.5m;180°;-12m/s;0m/s;0rad/s;$\mu(\theta\|c^{(1)})$]	2.2s	10s
T3	[-2.5m;120m;-90°;0m/s;-10m/s;0rad/s;$\mu(\theta\|c^{(1)})$]	4.2s	10s
T4	[2.5m;0m;90°;0m/s;10m/s;0rad/s;$\mu(\theta\|c^{(2)})$]	4.2s	10s

I, which is simple but sufficient. According to their shape sizes, they can represent cars, bicycles and trucks, respectively. The initial states of the four vehicles are given in Table II. The state vector includes the position of target center, heading orientation, velocities in two directions, yaw rate and customized radii of different classes. The total testing time is 10s with a sampling period of $\mathrm{T}=0.2\mathrm{s}$. The measurements are uniformly generated along the target contour with an observation noise whose covariance is given by $\bar{R}_{k,l}=0.1^2\mathrm{m}^2 I_2$. The clutters are assumed uniformly distributed in the whole surveillance area with a Poisson intensity of $\lambda=1$. The detection probability p_{D} and survival probability p_{S} are both 0.99. Although p_{D} and p_{S} are class conditioned in the proposed filter, the simulation doesn't make a distinction of them between classes for the sake of simplicity. The tuning parameters of UT are set as $\alpha=1,\beta=2,\kappa=2$. As we try to reduce the calculation, the number of star-convex basic radii is set by 8. The eight radii take the line connecting the rectangular center with four vertices and the midpoints of four edges. Nonetheless, that is well-informed for simple rectangular targets. The hyper-parameters for GP model are given by $\sigma_f=0.8\mathrm{m}$ and $\ell=\pi/2\mathrm{rad}$. According to (13), the customized variance of radius $\sigma_r^{2(i)}$ can be calculated by the standard deviation of length and width in Table I. Besides, the TPM of the classes and initial class probabilities are set by

$$T_{c|c'}=\begin{bmatrix}0.8 & 0.1 & 0.1\\ 0.1 & 0.8 & 0.1\\ 0.1 & 0.1 & 0.8\end{bmatrix}, \pi_0(c)=\begin{bmatrix}1/3 & 1/3 & 1/3\end{bmatrix}^{\top}. \quad (43)$$

The Fig. 3 shows the tracking and classification results of the four vehicles. The patterns of vehicles are drawn at intervals of 5 frames. As we see, the positions and shapes are correctly estimated. Meanwhile, the classification is also realized by the proposed CE-JTC-ET-PHD filter. Although only a simple and linear motion model is considered in this paper, other nonlinear dynamic models are also feasible due to the compatibility of the designed filter according to (32).

We also test the performance of proposed CE-JTC method compared with each single shape model with the same ET-

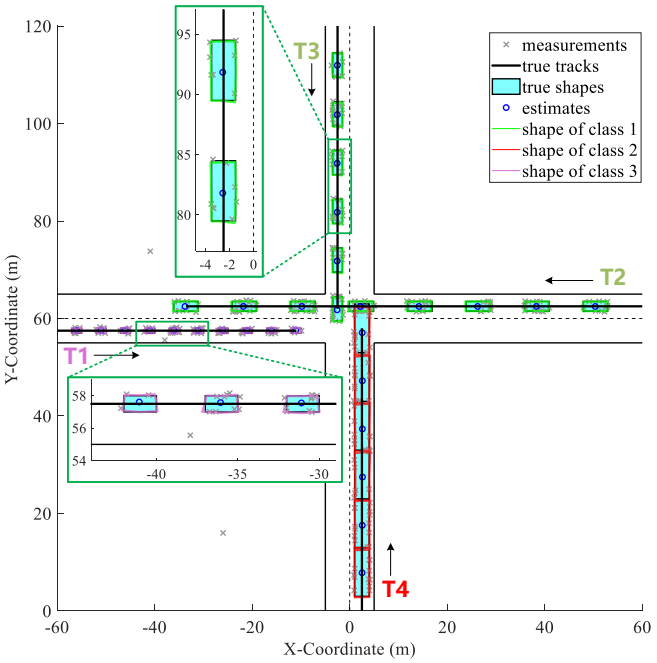

Fig. 3. The tracking and classification results of the four vehicles at an intersection. The black arrows represent the direction of movement. The blue circles are the estimated positions of targets. Different colors of the vehicle shapes represent the classification results.

PHD filter. The Fig. 4 shows the results for the interaction-over-union (IOU) [3] and optimal sub-patten assignment (OSPA) [32]. The IOU is computed as the interaction of estimated extent and the truth over the union of them. The OSPA employs a Euclidean 2-norm metric of target's center position. The curves in Fig. 4 show that the estimation accuracies of both targets' positions and shapes are better than only using a single vehicle model. The results demonstrate that the proposed algorithm not only successfully achieves JTC but also can improve the performance of shape and position estimation to a certain degree.

V. CONCLUSION

In this work, we propose a CE-JTC-ET-PHD filter based on customized GP models. In the proposed CE-JTC method, we leverage the shape size of vehicles as the indicator of classification and strengthen the role of class in the filtering. The numerical results confirm that the whole JTC method is practicable and effective in common traffic scenes. In future work, we will explore the use of more advanced filters like labeled RFS filters [33–35] and validate our proposed JTC method with real data collected from traffic sensors.

REFERENCES

[1] A. Scheel, S. Reuter, and K. Dietmayer, "Vehicle tracking using extended object methods: An approach for fusing radar and laser," in *Proc. Int. Conf. Robot. Automat.*, May 2017.

[2] J. Xiong, R. Bi, Y. Tian, X. Liu, and D. Wu, "Toward lightweight, privacy-preserving cooperative object classification for connected autonomous vehicles," *IEEE Internet Things J.*, vol. 9, no. 4, pp. 2787–2801, Feb. 2022.

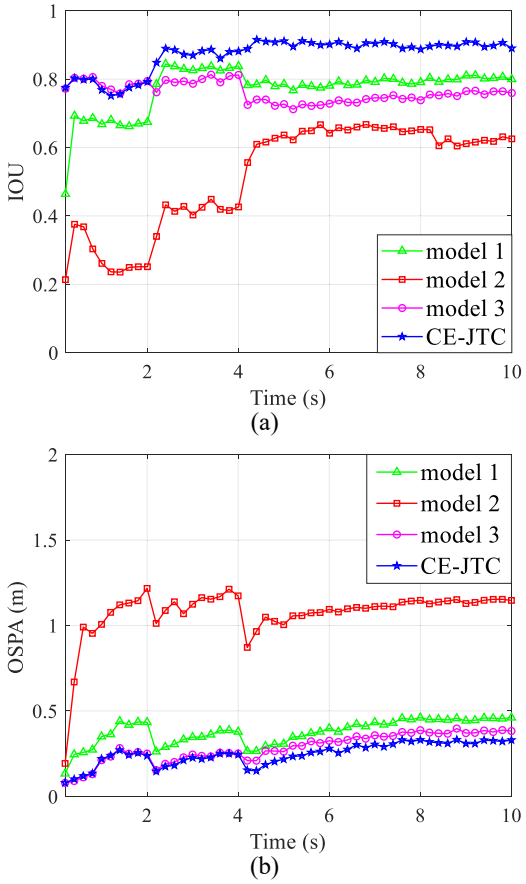

Fig. 4. The averaged IOU and OSPA curves of 100 Monte Carlo runs are generated from the proposed CE-JTC-ET-PHD filter and the ET-PHD filter only using the single shape model. The higher IOU means the better shape estimation. The lower OSPA means the more precise position estimation.

[3] K. Granström, C. Lundquist, and U. Orguner, "Tracking rectangular and elliptical extended targets using laser measurements," in *Proc. 14th Int. Conf. Inf. Fusion*, 2011, pp. 1–8.

[4] T. Kurien, "Framework for integrated tracking and identification of multiple targets," in *Proc. IEEE/AIAA 10th Digit. Avionics Syst. Conf.*, Los Angeles, CA, USA, Oct. 1991.

[5] R. Mahler, C. Rago, T. Zajic, S. Musick, and R. K. Mehra, "Joint tracking, pose estimation, and identification using HRRR data," in *Proc. SPIE Signal Process., Sensor Fusion, and Target Recognit. IX*, I. Kadar, Ed., vol. 4052, Aug. 2000, pp. 195–206.

[6] S. M. Herman, "Joint passive radar tracking and target classification using radar cross section," in *Proc. SPIE Signal and Data Process. of Small Targets*, vol. 5204, Dec. 2003, pp. 402–417.

[7] B. Ristic, N. Gordon, and A. Bessell, "On target classification using kinematic data," *Inform. Fusion*, vol. 5, no. 1, pp. 15–21, Mar. 2004.

[8] D. Angelova and L. Mihaylova, "Joint target tracking and classification with particle filtering and mixture kalman filtering using kinematic radar information," *Digital Signal Process.*, vol. 16, no. 2, pp. 180–204, Mar. 2006.

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

[10] K. Granström, S. Reuter, D. Meissner, and A. Scheel, "A multiple model PHD approach to tracking of cars under an assumed rectangular shape," in *Proc. 17th Int. Conf. Inf. Fusion*, 2014, pp. 1–8.

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

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

[13] K. Granström, M. Fatemi, and L. Svensson, "Gamma Gaussian inverse-Wishart Poisson multi-Bernoulli filter for extended target tracking," in *Proc. 19th Int. Conf. Inf. Fusion*, 2016, pp. 893–900.

[14] M. Beard, S. Reuter, K. Granström, B.-T. Vo, B.-N. Vo, and A. Scheel, "Multiple extended target tracking with labeled random finite sets," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1638–1653, Apr. 2016.

[15] M. Baum and U. D. Hanebeck, "Extended object tracking with random hypersurface models," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 1, pp. 149–159, Jan. 2014.

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

[17] K. Thormann, M. Baum, and J. Honer, "Extended target tracking using Gaussian processes with high-resolution automotive radar," in *Proc. 21st Int. Conf. Inf. Fusion*, Jul. 2018.

[18] T. Hirscher, A. Scheel, S. Reuter, and K. Dietmayer, "Multiple extended object tracking using Gaussian processes," in *Proc. 19th Int. Conf. Inf. Fusion*, 2016, pp. 868–875.

[19] M. Fröhle, K. Granström, and H. Wymeersch, "Decentralized Poisson Multi-Bernoulli filtering for vehicle tracking," *IEEE Access*, vol. 8, pp. 126 414–126 427, 2020.

[20] J. Yang, Y. Tian, H. Mao, Z. Zou, and W. Yi, "Distributed posterior fusion for vehicle tracking with Gaussian processes," in *Proc. 12th Int. Conf. Control, Automat. Inf. Sci.*, Hanoi, Vietnam, Nov. 2023, pp. 218–223.

[21] J. Lan and X. R. Li, "Joint tracking and classification of extended object using random matrix," in *Proc. 16th Int. Conf. Inf. Fusion*, Istanbul, Turkey, Jul. 2013, pp. 1550–1557.

[22] ——, "Joint tracking and classification of non-ellipsoidal extended object using random matrix," in *Proc. 17th Int. Conf. Inf. Fusion*, Salamanca, Spain, Jul. 2014, pp. 1–8.

[23] L. Wang, R. Zhan, S. Liu, J. Zhang, and Z. Zhuang, "Joint tracking and classification of multiple extended targets via the PHD filter and star-convex RHM," *Digital Signal Process.*, vol. 111, p. 102961, Apr. 2021.

[24] L. Wang and R. Zhan, "Joint detection, tracking and classification of multiple maneuvering star-convex extended targets," *IEEE Sensors J.*, pp. 1–1, 2024.

[25] B. Tuncer, M. Kumru, E. Ozkan, and A. A. Alatan, "Extended object tracking and shape classification," in *Proc. 21st Int. Conf. Inf. Fusion*, Jul. 2018.

[26] C. Magnant, A. Giremus, E. Grivel, L. Ratton, and B. Joseph, "Multi-target tracking using a PHD-based joint tracking and classification algorithm," in *Proc. IEEE Radar Conf.*, May 2016.

[27] R. Mahler, "PHD filters for nonstandard targets, i: Extended targets," in *Proc. 12th Int. Conf. Inf. Fusion*, 2009, pp. 915–921.

[28] K. Granström, C. Lundquist, and O. Orguner, "Extended target tracking using a Gaussian-mixture PHD filter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3268–3286, Oct. 2012.

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

[30] 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 Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.

[31] R. Van Der Merwe, *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. Oregon Health & Science University, 2004.

[32] D. Schuhmacher, B.-T. Vo, and B.-N. Vo, "A consistent metric for performance evaluation of multi-object filters," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3447–3457, Aug. 2008.

[33] B.-T. Vo and B.-N. Vo, "Labeled random finite sets and multi-object conjugate priors," *IEEE Trans. Signal Process.*, vol. 61, no. 13, pp. 3460–3475, Jul. 2013.

[34] B.-N. Vo, B.-T. Vo, and D. Phung, "Labeled random finite sets and the Bayes multi-target tracking filter," *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6554–6567, Dec. 2014.

[35] B.-N. Vo and B.-T. Vo, "An implementation of the multi-sensor generalized labeled multi-Bernoulli filter via Gibbs sampling," in *Proc. 20th Int. Conf. Inf. Fusion*, Xi'an, China, Jul. 2017, pp. 1–8.