Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Joint tracking and classification of extended targets with complex shapes*

Liping WANG, Ronghui ZHAN†‡, Yuan HUANG, Jun ZHANG, Zhaowen ZHUANG

National Key Laboratory of Science and Technology on Automatic Target Recognition,
National University of Defense Technology, *Changsha 410073*, *China*

†E-mail: zhanrh@nudt.edu.cn

Received Feb. 6, 2020; Revision accepted July 30, 2020; Crosschecked June 1, 2021

Abstract: This paper addresses the problem of joint tracking and classification (JTC) of a single extended target with a complex shape. To describe this complex shape, the spatial extent state is first modeled by star-convex shape via a random hypersurface model (RHM), and then used as feature information for target classification. The target state is modeled by two vectors to alleviate the influence of the high-dimensional state space and the severely nonlinear observation model on target state estimation, while the Euclidean distance metric of the normalized Fourier descriptors is applied to obtain the analytical solution of the updated class probability. Consequently, the resulting method is called the "JTC-RHM method." Besides, the proposed JTC-RHM is integrated into a Bernoulli filter framework to solve the JTC of a single extended target in the presence of detection uncertainty and clutter, resulting in a JTC-RHM-Ber filter. Specifically, the recursive expressions of this filter are derived. Simulations indicate that: (1) the proposed JTC-RHM method can classify the targets with complex shapes and similar sizes more correctly, compared with the JTC method based on the random matrix model; (2) the proposed method performs better in target state estimation than the star-convex RHM based extended target tracking method; (3) the proposed JTC-RHM-Ber filter has a promising performance in state detection and estimation, and can achieve target classification correctly.

Key words: Extended target; Fourier descriptors; Joint tracking and classification; Random hypersurface model; Bernoulli filter
https://doi.org/10.1631/FITEE.2000061 **CLC number:** TN953

1 Introduction

Target tracking and target classification are investigated widely in most surveillance systems. Target tracking aims to estimate the target state (i.e., an estimation), and target classification is about determining the target class (i.e., a decision). Traditionally, target tracking and target classification are usually tackled by a separate two-step strategy ("decision-then-estimation" or "estimation-then-decision"). Such an approach is generally suboptimal, as target tracking and target classification are mutually influential steps. Therefore, they should be processed jointly, i.e., as joint tracking and classification (JTC) (Ristic et al., 2004; Angelova and Mihaylova, 2006; Jiang et al., 2015; Cao et al., 2016, 2018; Magnant et al., 2018).

The process of JTC can make full use of information on both target estimation and decision, as well as improve the accuracy of both. For example, using class information, the class-dependent motion models can be selected for target tracking. In return, target tracking results can be used as features for target classification. For JTC, classification can be deduced from only kinematic characteristics (Ristic et al., 2004; Angelova and Mihaylova, 2006), or from complementary information provided by attribute sensors, such as electronic support measures (ESMs)

‡ Corresponding author
* Project supported by the National Natural Science Foundation of China (No. 61471370)
ORCID: Ronghui ZHAN, https://orcid.org/0000-0001-6799-620X

(Jiang et al., 2015; Cao et al., 2018).

In traditional target tracking, targets are assumed to be point targets. However, with sensor resolution improving constantly, this assumption no longer holds. A single target will occupy multiple resolution cells and generate multiple measurements resulting in an extended target. The extended target tracking (ETT) methods have been studied extensively with prolific results (Granström et al., 2012, 2017; Mihaylova et al., 2014; Beard et al., 2016; Eryildirim and Guldogan, 2016; de Freitas et al., 2019). For ETT, not only the kinematic state (position, velocity, acceleration, etc.) but also the spatial extent state (shape, size, and orientation) should be estimated. Currently, there are two main types of approach for modeling the extent state: simple geometric shape based approaches, and star-convex shape based approaches.

For the former, the extent state can be modeled as a stick (Gilholm and Salmond, 2005; Baum et al., 2012), a circle (Baum et al., 2010), a rectangle (Granström et al., 2014; Knill et al., 2016), or an ellipse (Angelova et al., 2013; Yang and Baum, 2016, 2017). The random matrix model (RMM) based methods (Koch, 2008; Feldmann et al., 2011; Lan and Li, 2016) have been studied widely by modeling the target shape as an ellipse. Star-convex shape based approaches are based on the random hypersurface model (RHM) (Baum and Hanebeck, 2014) or the Gaussian process model (Wahlström and Özkan, 2015; Hirscher et al., 2016). Simple geometric shape based approaches have proven to be effective in certain applications, e.g., tracking boats with elliptical shapes and cars with rectangular shapes. However, real-world targets may have much more complex shapes. In addition, the more accurate the spatial extent state estimations, the more favorable to achieve target classification.

Two main methods for tracking extended targets with complex shapes are the non-ellipsoid-based ETT method (Lan and Li, 2014; Granström et al., 2015) and the star-convex RHM based ETT method (Baum and Hanebeck, 2014). In the non-ellipsoid-based ETT method, it is assumed that the number of sub-ellipses is known, while this method is not universal for different types of target; however, the star-convex RHM based ETT method can track targets with diverse shapes.

Although the latter provides an effective way for tracking of an extended target with a complex shape, some problems remain to be solved. First, the prior shape of the target is initialized as a circle regardless of the true target shape. Although this shape can be approached gradually by recursive filtering, the real-time requirements of target tracking are not met. Second, the target state is modeled as a high-dimensional vector with the observation model being highly nonlinear. Thus, the performance of the star-convex RHM based ETT method will degrade when the target performs a maneuver.

Compared with the JTC of point targets, that of extended targets aims to estimate kinematic and extent states, and to decide the target class simultaneously. Since different classes of extended targets have unique spatial extent forms, the extent state can be used as a feature for target classification (Lan and Li, 2013; Hu et al., 2018; Sun et al., 2018). A JTC method based on RMM (JTC-RMM) that integrates the class-dependent extent information into the RMM-based ETT method was proposed by Lan and Li (2013). The extent state in the JTC-RMM method is modeled as an elliptical shape using a symmetric positive definite (SPD) matrix. Although the JTC-RMM method provides a novel mean for the JTC of extended targets, the drawback is that the performance of this method degrades when targets have similar sizes, as the extent state is simply modeled as an ellipse shape, with only the size feature used for target classification.

To overcome the drawback of the JTC-RMM method, and to use the extent state as the classification feature (both shape and size information) more effectively, we adopt the star-convex RHM in this study to model the extent state, and use the extent state as a classification feature. We propose a JTC method (called the "JTC-RHM method") for extended targets with a complex shape. The proposed method models the target state with two vectors, and can solve the aforementioned two problems with the conventional star-convex RHM based ETT method. The main contributions of this study are summarized as follows:

1. The proposed JTC-RHM method can classify targets correctly when targets have similar sizes but with different shapes using the class-dependent and received measurements.

2. Under the nonlinear observation model, it is difficult to obtain an analytical solution for the

calculation of the class probability of target classes as Eq. (24) of the JTC-RMM method in Lan and Li (2013). Fourier descriptors are used to tackle this problem. Particularly, to alleviate the influence of the severe nonlinear observation model and high-dimensional state vector on target tracking in the conventional star-convex RHM based ETT method, the target state is modeled using two vectors and target estimation is thus improved.

3. The proposed JTC-RHM method is integrated into the Bernoulli filter framework for the JTC of a single extended target in the presence of detection uncertainty and clutter, resulting in a JTC-RHM-Ber filter. Specifically, the expressions for the JTC-RHM-Ber filter recursions are derived and presented.

2 Background

In this section, the star-convex RHM based ETT method is presented first, and some basic models and formulations used in this study are introduced. We then present the Bayesian inference based JTC principle.

2.1 Star-convex RHM based ETT method

For the star-convex RHM based ETT method, the target state at time k is modeled by a high-dimensional vector, denoted as $\boldsymbol{x}_k$. State $\boldsymbol{x}_k$ consists of two parts: kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ and extent state $\boldsymbol{x}_k^{\mathrm{e}}$, i.e., $\boldsymbol{x}_k \triangleq \left[(\boldsymbol{x}_k^{\mathrm{k}})^{\mathrm{T}}, (\boldsymbol{x}_k^{\mathrm{e}})^{\mathrm{T}}\right]^{\mathrm{T}}$. The kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ used in this study is a four-dimensional vector with a form of $\boldsymbol{x}_k^{\mathrm{k}} \triangleq [\boldsymbol{x}_k^{\mathrm{p}}, \boldsymbol{x}_k^{\mathrm{v}}]^{\mathrm{T}}$, where $\boldsymbol{x}^{\mathrm{p}} \triangleq (x, y)$ and $\boldsymbol{x}_k^{\mathrm{v}} \triangleq (\dot{x}, \dot{y})$ are the position and velocity of the target centroid, respectively. The dynamic evolution of kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ is described by a linear Gaussian dynamic model as

$$\boldsymbol{x}_{k+1}^{\mathrm{k}} = \boldsymbol{F}_k^{\mathrm{k}} \boldsymbol{x}_k^{\mathrm{k}} + \boldsymbol{w}_k^{\mathrm{k}}, \tag{1}$$

where $\boldsymbol{F}_k^{\mathrm{k}}$ is the evolution matrix and $\boldsymbol{w}_k^{\mathrm{k}}$ is a zero-mean Gaussian process noise with covariance matrix $\boldsymbol{Q}_k^{\mathrm{k}}$.

For an extended target, there are multiple measurements at each sampling instant. At time k, a set of received measurements is denoted as $\mathcal{Z}_k = \{\boldsymbol{z}_{k,l}\}_{l=1}^{n_{k,z}}$, where $n_{k,z}$ is the number of measurements at time k. It is assumed that each measurement $\boldsymbol{z}_{k,l}$ is generated using a measurement source $\boldsymbol{y}_{k,l}$ from the surface of an extended target:

$$\boldsymbol{z}_{k,l} = \boldsymbol{y}_{k,l} + \boldsymbol{v}_{k,l}, \tag{2}$$

where $\boldsymbol{v}_{k,l}$ is a zero-mean Gaussian observation noise with covariance matrix $\boldsymbol{R}_k$.

As shown in Eq. (2), the locations of measurement sources are unknown. For the star-convex RHM based ETT method, the shape of the extent state is described by the star-convex shape, and the locations of measurement sources are modeled using RHM based on the assumption that measurement sources are on a scaled boundary of the true shape of the target.

Definition 1 (Star-convex shape) If all points on the line connecting an arbitrary point in set $\mathcal{S}_{\boldsymbol{x}_k}$ and the centroid still belong to this set, the shape formed by set $\mathcal{S}_{\boldsymbol{x}_k}$ is then a star-convex shape.

A radial function $r(\phi_k)$ that describes the distance from the target centroid to the boundary point with angle ϕ_k can be used to represent the star-convex shape. Therefore, the star-convex shape can be expressed as

$$\mathcal{S}_{\boldsymbol{x}_k} \triangleq r(\phi_k)\boldsymbol{e}_{\phi_k} + \boldsymbol{x}_k^{\mathrm{p}}, \tag{3}$$

where $\phi_k \in [0, 2\pi)$ is the angle between the positive direction of the x axis and the line connecting the centroid and a boundary point. Variable $\boldsymbol{e}_{\phi_k} = [\cos\phi_k, \sin\phi_k]^{\mathrm{T}}$ is a two-dimensional (2D) vector with an angle ϕ_k.

The radial function can be expanded by the Fourier series as

$$r(\phi_k) = a_{k,0} + \sum_{n=1,2,\cdots,n_{\mathrm{F}}} [a_{k,n}\cos(n\phi_k) + b_{k,n}\sin(n\phi_k)]. \tag{4}$$

Therefore, $\mathcal{S}_{\boldsymbol{x}_k}$ can be parameterized by a vector composed of these Fourier coefficients $\boldsymbol{x}_k^{\mathrm{e}} \triangleq \left[a_{k,0}, a_{k,1}, b_{k,1}, a_{k,2}, b_{k,2}, \cdots, a_{k,n_{\mathrm{F}}}, b_{k,n_{\mathrm{F}}}\right]^{\mathrm{T}}$, i.e.,

$$\mathcal{S}_{\boldsymbol{x}_k} \triangleq \boldsymbol{r}_{\phi_k} \boldsymbol{x}_k^{\mathrm{e}} \boldsymbol{e}_{\phi_k} + \boldsymbol{x}_k^{\mathrm{p}}, \tag{5}$$

where $\boldsymbol{r}_{\phi_k} \triangleq \left[1, \cos\phi_k, \sin\phi_k, \cos(2\phi_k), \sin(2\phi_k), \cdots, \cos(n_{\mathrm{F}}\phi_k), \sin(n_{\mathrm{F}}\phi_k)\right]$, and the dimension of target state $\boldsymbol{x}_k$ is $n_{\mathrm{x}}=5+2n_{\mathrm{F}}$.

The location of the measurement source can be written as

$$\boldsymbol{y}_{k,l} = s_{k,l} \boldsymbol{r}_{\phi_{k,l}} \boldsymbol{x}_k^{\mathrm{e}} \boldsymbol{e}_{\phi_{k,l}} + \boldsymbol{x}_k^{\mathrm{p}}, \tag{6}$$

where $s_{k,l} \in [0, 1]$ is the scaling factor.

Combining Eqs. (2) and (6), the observation model of the convex-star RHM based ETT method becomes

$$\boldsymbol{z}_{k,l} = s_{k,l} \boldsymbol{r}_{\phi_{k,l}} \boldsymbol{x}_k^{\mathrm{e}} \boldsymbol{e}_{\phi_{k,l}} + \boldsymbol{x}_k^{\mathrm{p}} + \boldsymbol{v}_{k,l}. \tag{7}$$

Angle $\phi_{k,l}$ is unknown and can be estimated by the angle $\hat{\phi}_{k,l}$ between the line from a received position measurement $\boldsymbol{z}_{k,l}$ to the estimated centroid and the positive direction of the x axis. Algebraic manipulations are conducted to reduce the effect of $\hat{\phi}_{k,l}$ on the state estimation, with a resulting new observation model of

$$\begin{aligned} 0 = {} & s_{k,l}^2 \| \boldsymbol{x}_k^{\mathrm{e}} \boldsymbol{r}_{\phi_{k,l}} \|^2 + 2 s_{k,l} \boldsymbol{x}_k^{\mathrm{e}} \boldsymbol{r}_{\phi_{k,l}} \boldsymbol{e}_{\phi_{k,l}} \boldsymbol{v}_{k,l} + \| \boldsymbol{v}_{k,l} \|^2 \\ & - \| \boldsymbol{z}_{k,l} - \boldsymbol{x}_k^{\mathrm{p}} \|^2 . \end{aligned} \tag{8}$$

The right part of Eq. (8) can be abbreviated as $h^*(\boldsymbol{x}_k, \boldsymbol{v}_{k,l}, s_{k,l}, \boldsymbol{z}_{k,l})$, where $h^*(\cdot)$ maps target state $\boldsymbol{x}_k$, observation noise $\boldsymbol{v}_{k,l}$, scaling factor $s_{k,l}$, and measurement $\boldsymbol{z}_{k,l}$ into the pseudo-measurement 0. We use Eq. (8) as the observation model in this study.

2.2 Bayesian inference based JTC

It is assumed that there are n_c different target classes in the surveillance area, and each target has a class label $c \in \mathcal{C}$, where $\mathcal{C} \triangleq \{1, 2, \cdots, n_c\}$ is the class label set. Parameter $\mu^c \triangleq p(c)$ is used to denote the class probability. According to Bayesian theory, JTC aims to obtain the probability density-mass function (PDMF) $p(\boldsymbol{x}_k, c|\mathcal{Z}^k)$ based on measurements $\mathcal{Z}^k$ at time k, and the prior PDMF $p(\boldsymbol{x}_{k-1}, c|\mathcal{Z}^{k-1})$ at time $k-1$. Using the conditional probability equation, $p(\boldsymbol{x}_k, c|\mathcal{Z}^k)$ can be written as

$$p(\boldsymbol{x}_k, c \mid \mathcal{Z}^k) = p(\boldsymbol{x}_k \mid c, \mathcal{Z}^k) p(c \mid \mathcal{Z}^k), \tag{9}$$

where $\mathcal{Z}^k=(\mathcal{Z}_1, \mathcal{Z}_2, \cdots, \mathcal{Z}_k)$ denotes all measurements up to time k.

For target tracking, the class-dependent probability density function (PDF) $p(\boldsymbol{x}_k, c|\mathcal{Z}^k)$ can be obtained by

$$p(\boldsymbol{x}_k \mid c, \mathcal{Z}^k) = \frac{p(\mathcal{Z}_k \mid \boldsymbol{x}_k, c) p(\boldsymbol{x}_k \mid c, \mathcal{Z}^{k-1})}{p(\mathcal{Z}_k \mid c, \mathcal{Z}^{k-1})}, \tag{10}$$

where $p(\mathcal{Z}_k| c, \mathcal{Z}^{k-1}) = \int p(\mathcal{Z}_k| \boldsymbol{x}_k, c) p(\boldsymbol{x}_k| c, \mathcal{Z}^{k-1}) \mathrm{d}\boldsymbol{x}_k$ is the normalizing factor.

For target classification, the probability mass function (PMF) $\mu_k^c \triangleq p(c|\mathcal{Z}^k)$ can be obtained by

$$\mu_k^c = \frac{p(\mathcal{Z}_k \mid c, \mathcal{Z}^{k-1}) p(c|\mathcal{Z}^{k-1})}{p(\mathcal{Z}_k \mid \mathcal{Z}^{k-1})}, \tag{11}$$

where $p(\mathcal{Z}_k \mid \mathcal{Z}^{k-1}) = \sum_{c=1}^{n_c} p(\mathcal{Z}_k \mid c, \mathcal{Z}^{k-1}) p(c|\mathcal{Z}^{k-1})$ is the normalizing factor.

3 The proposed JTC-RHM method

The JTC-RMM method uses only the size information contained in the extent state for target classification. When targets have similar sizes, the JTC-RMM method cannot correctly classify these targets. We proposed an RHM based JTC (JTC-RHM) method to overcome this drawback, where the star-convex RHM is adopted to model the extent state to make full use of extent information. Different from the JTC-RMM method, we use the class-dependent feature points (rather than the SPD matrix) as the prior class-dependent measurement information, and incorporate this information into the star-convex RHM based ETT method. Due to the high nonlinearity of the observation model (8), and the fact that the state is modeled as a high-dimensional vector, it is hard to obtain the analytical solution of the updated

target class probability, and thus the performance of state estimation will be deteriorated greatly. Therefore, we model the state using two vectors and use the Euclidean distance metric of the normalized Fourier descriptors to calculate the updated class probability.

3.1 Prior class-dependent measurement information

The JTC method uses not only the measurements received through sensors, but also the prior class-dependent information. For a known class of an extended target, the positions of the feature points on the target contour can be acquired in advance. According to the principle of the star-convex RHM based ETT method, we can obtain an accurate shape estimate with these feature points by considering these points as the sources of measurement. The reason is that there is no need to consider the distribution of a scaling factor. Fig. 1 shows the estimation results of the extent state using two types of measurement. One type is generated by 24 feature points (denoted by red points), and others are 1000 measurements distributed randomly on the target surface (denoted by black points). The estimated extent state using the 24 feature points is similar to that using the 1000 random measurements.

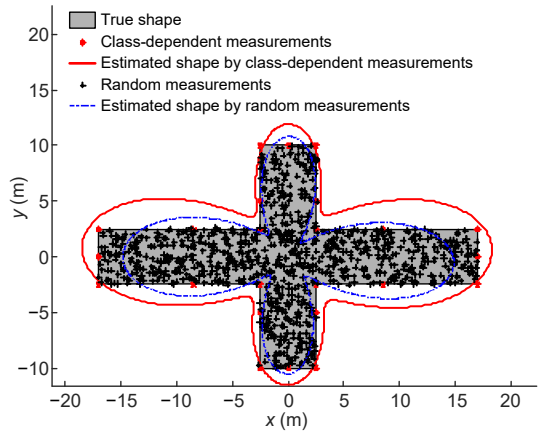

Fig. 1 Estimation results of the extent state using two kinds of measurements (References to color refer to the online version of this figure)

In this study, these feature points can be used as the prior information of the target class, which can be represented as $\mathcal{Z}^c \triangleq \{z^{c,i}\}_{i=1}^{n^{c,\mathrm{p}}}$, $c=1,2,\cdots,n_c$, where $n^{c,\mathrm{p}}$ is the number of feature points of the target class c. Using the origin of the coordinates as the reference point, we can obtain the position of prior class-dependent measurement information and prior class-dependent extent state $\boldsymbol{x}_0^{\mathrm{e},c}$. As the target moves, the prior class-dependent measurement information will change. Therefore, measurements at time k can be written as $\underline{\mathcal{Z}}_k \triangleq \{\mathcal{Z}_k, \{\mathcal{Z}_k^c\}_{c=1}^{n_c}\}$.

For target classification, the class probability $\tilde{\mu}_k^c$ of the proposed JTC-RHM method is calculated by

$$\tilde{\mu}_k^c \triangleq p(c \mid \underline{\mathcal{Z}}^k) = \frac{p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid c, \underline{\mathcal{Z}}^{k-1}) p(c \mid \underline{\mathcal{Z}}^{k-1})}{p(\underline{\mathcal{Z}}_k \mid \underline{\mathcal{Z}}^{k-1})}, \quad (12)$$

where $p(\underline{\mathcal{Z}}_k \mid \underline{\mathcal{Z}}^{k-1}) = \sum_{c=1}^{n_c} p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid c, \underline{\mathcal{Z}}^{k-1}) p(c \mid \underline{\mathcal{Z}}^{k-1})$ is the normalizing factor.

3.2 Calculation of the updated class probability

Due to the high dimensionality of state vectors and the nonlinearity of the observation model, it is difficult to derive the analytical solution of Eq. (12). For JTC of an extended target, the extent state can be used as the classification feature. Therefore, the proposed JTC-RHM method uses $\underline{\mathcal{Z}}_k \triangleq \{\mathcal{Z}_k, \{\mathcal{Z}_k^c\}_{c=1}^{n_c}\}$ to estimate extent state $\boldsymbol{x}_k^{\mathrm{e}}$. By modeling the relationship between extent state $\boldsymbol{x}_k^{\mathrm{e}}$ and the prior class-dependent extent state $\boldsymbol{x}_0^{\mathrm{e},c}$, the proposed JTC-RHM method uses this relationship to substitute the calculation of $p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid c, \underline{\mathcal{Z}}^{k-1})$.

If the received measurements $\mathcal{Z}^k$ are originated by target class c, the difference between the updated extent state $\boldsymbol{x}_k^{\mathrm{e}}$ and the prior distribution $\boldsymbol{x}_0^{\mathrm{e},c}$ is minimal. In a 2D plane, the shape of an extended target can be described using a closed curve. The Fourier descriptors are the Fourier transform coefficients of the boundary closure curve (Zhao and Belkasim, 2012). Therefore, we use the Euclidean distance of Fourier descriptors to quantitatively describe the difference between the shapes represented by $\boldsymbol{x}_k^{\mathrm{e}}$ and $\boldsymbol{x}_0^{\mathrm{e},c}$. The Fourier descriptors can be obtained in two steps:

1. The first step is to obtain a complex form using coordinates (x_i, y_i), i.e., $u_i = x_i + \mathrm{j}y_i$, where $i=0, 1,\ldots, n_{\mathrm{D}}-1$ and $\mathrm{j}=\sqrt{-1}$. Here, (x_i, y_i) and n_{D} denote the

point along the boundary closure curve and the number of these points, respectively.

2. The second step is to apply the one-dimensional (1D) discrete Fourier transform (DFT) as

$$f_m=\mathcal{F}(u_i)=\sum_{i=0}^{n_{\mathrm{D}}-1}\left[u_i\exp\left(-\mathrm{j}\frac{2\pi im}{n_{\mathrm{D}}}\right)\right], \tag{13}$$

where m=0, 1,…, n_{D}−1, $\mathcal{F}(\cdot)$ represents 1D DFT, and f_m denotes the Fourier descriptors.

To keep the Fourier descriptors invariant to translation and rotation, f_m should be normalized. Thus, the normalized Fourier descriptors are written as

$$\tilde{f}_{\tilde{m}}=\frac{\|f_{g'}\|}{\|f_1\|},\ \tilde{m}=1,2,\cdots,n_{\mathrm{D}}-1, \tag{14}$$

where $\|\cdot\|$ denotes the modular operation.

Therefore, the Euclidean distance of two normalized Fourier descriptors $\tilde{f}_{\tilde{m}}^1$ and $\tilde{f}_{\tilde{m}}^2$ are used to describe the shape similarity of two targets:

$$\eta=\sqrt{\sum_{\tilde{m}=1}^{n_{\mathrm{D}}-1}\|\tilde{f}_{\tilde{m}}^1-\tilde{f}_{\tilde{m}}^2\|^2}. \tag{15}$$

If distance η=0, then two targets have the same contour. The larger the distance η, the more significant the difference between the contours of two targets.

If the received measurements $\mathcal{Z}^k$ are generated using target class c, then $p(\mathcal{Z}_k,\mathcal{Z}_k^c\,|\,c,\underline{\mathcal{Z}}^{k-1})$ of target class c in Eq. (12) should be larger than those of other classes. Hence, the proposed JTC-RHM method uses the Euclidean distance of the normalized Fourier descriptors to evaluate the relationship between extent state $\boldsymbol{x}_k^{\mathrm{e}}$ and the prior class-dependent extent state $\boldsymbol{x}_0^{\mathrm{e},c}$, and thus the calculation of $p(\mathcal{Z}_k,\mathcal{Z}_k^c|\,c,\underline{\mathcal{Z}}^{k-1})$ is substituted by

$$\tilde{\eta}_k^c=\exp\left(-\frac{1}{\sigma}\sqrt{\sum_{\tilde{m}=1}^{n_{\mathrm{D}}-1}\|\tilde{f}_{\tilde{m},0}^c-\tilde{f}_{\tilde{m},k}^c\|^2}\right), \tag{16}$$

where $\tilde{f}_{\tilde{m},0}^c$ and $\tilde{f}_{\tilde{m},k}^c$ are the normalized Fourier descriptors of the prior class-dependent extent state $\boldsymbol{x}_0^{\mathrm{e},c}$ and estimated extent state $\boldsymbol{x}_k^{\mathrm{e}}$ obtained by $p(\mathcal{Z}_k,\mathcal{Z}_k^c\,|\,c,\underline{\mathcal{Z}}^{k-1})$ at time k, respectively, parameter σ is the weight parameter set at σ=10^{-2}, whereas $\tilde{\eta}_k^c$ is used to calculate the updated probability of the target class.

Accordingly, Eq. (12) can be rewritten as

$$\tilde{\mu}_k^c=\frac{\tilde{\eta}_k^c\tilde{\mu}_{k-1}^c}{\sum_{c=1}^{n_{\mathrm{c}}}\tilde{\eta}_k^c\tilde{\mu}_{k-1}^c}. \tag{17}$$

3.3 Main steps of the JTC-RHM method

The observation model (7) can be rewritten as

$$\begin{aligned}\boldsymbol{z}_{k,l}&=s_{k,l}\boldsymbol{r}_{\phi_{k,l}}\boldsymbol{x}_k^{\mathrm{e}}\boldsymbol{e}_{\phi_{k,l}}+\boldsymbol{x}_k^{\mathrm{p}}+\boldsymbol{v}_{k,l}\\&=s_{k,l}\boldsymbol{r}_{\phi_{k,l}}\boldsymbol{x}_k^{\mathrm{e}}\boldsymbol{e}_{\phi_{k,l}}+\boldsymbol{H}\boldsymbol{x}_k^{\mathrm{k}}+\boldsymbol{v}_{k,l}\\&=\boldsymbol{H}\boldsymbol{x}_k^{\mathrm{k}}+\tilde{\boldsymbol{v}}_{k,l},\end{aligned} \tag{18}$$

where $\boldsymbol{H}$=[$\boldsymbol{I}_d$, $\boldsymbol{0}_{d\times d}$], $\boldsymbol{I}_d$ and $\boldsymbol{0}_{d\times d}$ are the unit matrix and an all-zero matrix with a size of $d\times d$, respectively. Here, d=2 is the dimension of the target motion space, while $\tilde{\boldsymbol{v}}_{k,l}$ is a zero-mean Gaussian measurement noise with covariance matrix $\tilde{\boldsymbol{R}}_k$.

On one hand, the observation model for kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ is linearly conditioned on the received measurements $\mathcal{Z}_k$ (model (18)). On the other hand, state vector $\boldsymbol{x}_k$ is high-dimensional, and this will lead to a poor state estimation. For the state estimation of the proposed JTC-RHM, we are concerned about the estimation of density $p(\boldsymbol{x}_k^{\mathrm{k}},\boldsymbol{x}_k^{\mathrm{e}}\,|\,c,\underline{\mathcal{Z}}^k)$. Following the concept of conditional distribution, posterior class-dependent density $p(\boldsymbol{x}_k^{\mathrm{k}},\boldsymbol{x}_k^{\mathrm{e}}\,|\,c,\underline{\mathcal{Z}}^k)$ can be decomposed into two conditional densities as

$$p(\boldsymbol{x}_k^{\mathrm{k}},\boldsymbol{x}_k^{\mathrm{e}}\,|\,c,\underline{\mathcal{Z}}^k)=p(\boldsymbol{x}_k^{\mathrm{e}}\,|\,\boldsymbol{x}_k^{\mathrm{k}},c,\underline{\mathcal{Z}}^k)\,p(\boldsymbol{x}_k^{\mathrm{k}}\,|\,c,\underline{\mathcal{Z}}^k). \tag{19}$$

Furthermore, we use the same dynamic evolution for all targets. Thus, kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ depends on only the received measurements $\mathcal{Z}_k$; i.e., $p(\boldsymbol{x}_k^{\mathrm{k}}\,|\,c,\underline{\mathcal{Z}}^k)=p(\boldsymbol{x}_k^{\mathrm{k}}\,|\,\mathcal{Z}^k)$. Accordingly, Eq. (19) can

be written as

$$p(\boldsymbol{x}_k^{\mathrm{k}}, \boldsymbol{x}_k^{\mathrm{e}} \mid c, \underline{\mathcal{Z}}^k) = p(\boldsymbol{x}_k^{\mathrm{e}} \mid \boldsymbol{x}_k^{\mathrm{k}}, c, \underline{\mathcal{Z}}^k) p(\boldsymbol{x}_k^{\mathrm{k}} \mid \mathcal{Z}^k). \quad (20)$$

Therefore, the target state is not modeled as $\boldsymbol{x}_k \triangleq \left[(\boldsymbol{x}_k^{\mathrm{k}})^{\mathrm{T}}, (\boldsymbol{x}_k^{\mathrm{e}})^{\mathrm{T}}\right]^{\mathrm{T}}$ in the proposed JTC-RHM method as opposed to the conventional star-convex RHM based ETT method, but it is modeled as $\underline{\boldsymbol{x}}_k \triangleq (\boldsymbol{x}_k^{\mathrm{k}}, \boldsymbol{x}_k^{\mathrm{e}})$.

According to Eq. (20), the Bayesian estimator for target tracking in the proposed JTC-RHM method is

$$\begin{aligned} & p(\underline{\boldsymbol{x}}_k \mid c, \underline{\mathcal{Z}}^k) \\ &= \frac{p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid \underline{\boldsymbol{x}}_k, c, \underline{\mathcal{Z}}^{k-1}) p(\boldsymbol{x}_k^{\mathrm{e}} \mid \boldsymbol{x}_k^{\mathrm{k}}, c, \underline{\mathcal{Z}}^{k-1})}{p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid \boldsymbol{x}_k^{\mathrm{k}}, c, \underline{\mathcal{Z}}^{k-1})} \\ &\quad \cdot \frac{p(\mathcal{Z}_k \mid \boldsymbol{x}_k^{\mathrm{k}}) p(\boldsymbol{x}_k^{\mathrm{k}} \mid \mathcal{Z}^{k-1})}{p(\mathcal{Z}_k \mid \mathcal{Z}^{k-1})}, \end{aligned} \quad (21)$$

with

$$\begin{aligned} & p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid \boldsymbol{x}_k^{\mathrm{k}}, c, \underline{\mathcal{Z}}^{k-1}) \\ &= \int p(\mathcal{Z}_k, \mathcal{Z}_k^c \mid \underline{\boldsymbol{x}}_k, c, \underline{\mathcal{Z}}^{k-1}) p(\boldsymbol{x}_k^{\mathrm{e}} \mid \boldsymbol{x}_k^{\mathrm{k}}, c, \underline{\mathcal{Z}}^{k-1}) \mathrm{d}\boldsymbol{x}_k^{\mathrm{e}}, \end{aligned} \quad (22)$$

$$p(\mathcal{Z}_k \mid \mathcal{Z}^{k-1}) = \int p(\mathcal{Z}_k \mid \boldsymbol{x}_k^{\mathrm{k}}) p(\boldsymbol{x}_k^{\mathrm{k}} \mid \mathcal{Z}^{k-1}) \, \mathrm{d}\boldsymbol{x}_k^{\mathrm{k}}. \quad (23)$$

Remark 1 Such state modeling and decomposition will result in an analytical expression for the estimation of kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$, and a tractable approximation estimation of extent state $\boldsymbol{x}_k^{\mathrm{e}}$. For the estimation of kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$, the linear observation model (18) is used. To ensure the equivalence of models (7) and (18), $\tilde{\boldsymbol{R}}_k$ is assumed to be dependent on the target extension and observation error. Here we adopt the model proposed in Lan and Li (2016) to describe the covariance matrix of observation noise $\tilde{\boldsymbol{R}}_k \triangleq \boldsymbol{B}_k \boldsymbol{X}_k \boldsymbol{B}_k$, where $\boldsymbol{X}_k$ and $\boldsymbol{B}_k$ are matrices with size $d \times d$. Matrix $\boldsymbol{X}_k$ follows the inverse Wishart distribution $\mathcal{IW}(\boldsymbol{X}_k; v_k, \boldsymbol{V}_k)$, and is used to describe the target extension. Matrix $\boldsymbol{B}_k$ describes the distortion of the target extension. Besides, target extension is considered for modeling the process noise, whereas the model in Lan and Li (2016) is used to describe the process noise covariance matrix $\boldsymbol{Q}_k^{\mathrm{k}} \triangleq \tilde{\boldsymbol{Q}}_k^{\mathrm{k}} \otimes \boldsymbol{X}_k$, where $\tilde{\boldsymbol{Q}}_k^{\mathrm{k}}$ is the covariance matrix of process noise in 1D space, and $\otimes$ denotes the operation of Kronecker product. Specifically, we use a Kalman-like filter proposed by Lan and Li (2016) to obtain the kinematic state estimation. For the estimation of extent state $\boldsymbol{x}_k^{\mathrm{e}}$, the class-dependent measurements $\mathcal{Z}_k^c$ and the received measurements $\mathcal{Z}_k$ are processed sequentially by the unscented Kalman filter (UKF) with observation model (8).

According to the dynamic evolution model (1) of kinematic state, the observation model (8) of the star-convex RHM based ETT method, and the linear observation model (18) of kinematic state, the posterior class-dependent target state of the proposed JTC-RHM method follows a Gaussian–Gaussian distribution at time $k-1$:

$$\begin{aligned} & p(\underline{\boldsymbol{x}}_{k-1} \mid c, \underline{\mathcal{Z}}^{k-1}) \\ &= p(\boldsymbol{x}_{k-1}^{\mathrm{e}} \mid \boldsymbol{x}_{k-1}^{\mathrm{k}}, c, \underline{\mathcal{Z}}^{k-1}) p(\boldsymbol{x}_{k-1}^{\mathrm{k}} \mid \mathcal{Z}^{k-1}) \\ &= \mathcal{N}(\boldsymbol{x}_{k-1}^{\mathrm{e}}; \boldsymbol{m}_{k-1}^{\mathrm{e},c}, \boldsymbol{P}_{k-1}^{\mathrm{e},c}) \mathcal{N}(\boldsymbol{x}_{k-1}^{\mathrm{k}}; \boldsymbol{m}_{k-1}^{\mathrm{k}}, \boldsymbol{P}_{k-1}^{\mathrm{k}}), \end{aligned} \quad (24)$$

where $\mathcal{N}(\boldsymbol{x}; \boldsymbol{m}, \boldsymbol{P})$ denotes the Gaussian PDF of a random vector $\boldsymbol{x}$ with mean $\boldsymbol{m}$ and covariance matrix $\boldsymbol{P}$. Taking target extension into consideration in extent state estimation, $\boldsymbol{P}_{k-1}^{\mathrm{k}}$ is modeled as $\boldsymbol{P}_{k-1}^{\mathrm{k}} \triangleq \tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k}} \otimes \boldsymbol{X}_{k-1}$, where $\tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k}}$ is the covariance matrix of the target kinematic state in the 1D space (Lan and Li, 2016).

The recursive process of the JTC-RHM method includes two steps: prediction and update.

Proposition 1 Assuming that the posterior class-dependent target state $\underline{\boldsymbol{x}}_{k-1}$ follows a Gaussian–Gaussian distribution as in Eq. (24) at time $k-1$, under the dynamic evolution model of the kinematic state as in model (1), the predicted target state $\underline{\boldsymbol{x}}_k$ also follows a Gaussian–Gaussian distribution:

$$\begin{aligned} & p(\underline{\boldsymbol{x}}_k \mid c, \underline{\mathcal{Z}}^{k-1}) \\ &= \mathcal{N}(\boldsymbol{x}_k^{\mathrm{e}}; \boldsymbol{m}_{k|k-1}^{\mathrm{e},c}, \boldsymbol{P}_{k|k-1}^{\mathrm{e},c}) \mathcal{N}(\boldsymbol{x}_k^{\mathrm{k}}; \boldsymbol{m}_{k|k-1}^{\mathrm{k}}, \boldsymbol{P}_{k|k-1}^{\mathrm{k}}). \end{aligned} \quad (25)$$

We use the Kalman-like filter to predict the kinematic state. For covariance matrix $\boldsymbol{P}_{k|k-1}^{\mathrm{k}}$, we focus on only its 1D representation for recursive

filtering. Therefore, the spatial PDF of the predicted kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ is $\mathcal{N}(\boldsymbol{x}_k^{\mathrm{k}};\boldsymbol{m}_{k|k-1}^{\mathrm{k}},\boldsymbol{P}_{k|k-1}^{\mathrm{k}})$ with

$$\boldsymbol{m}_{k|k-1}^{\mathrm{k}}=\boldsymbol{F}_k^{\mathrm{k}}\boldsymbol{m}_{k-1}^{\mathrm{k}}, \tag{26}$$

$$\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k}}=\tilde{\boldsymbol{F}}_k^{\mathrm{k}}\tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k}}(\tilde{\boldsymbol{F}}_k^{\mathrm{k}})^{\mathrm{T}}+\tilde{\boldsymbol{Q}}_k^{\mathrm{k}}, \tag{27}$$

where $\tilde{\boldsymbol{F}}_k^{\mathrm{k}}$ is the 1D representation of evolution matrix $\boldsymbol{F}_k^{\mathrm{k}}$.

Remark 2 In our proposed JTC-RHM method, only the kinematic state is handled recursively. Accordingly, we predict only the kinematic state. To incorporate prior class-dependent measurement information into the star-convex RHM based ETT method, and to ensure that the estimation results of these measurements are not affected by the received measurements up to time k, the extent state is processed sequentially in the update step by only the prior class-dependent measurement information and received measurements at time k. Thus, the mean and covariance matrix of extent state $\boldsymbol{x}_k^{\mathrm{e}}$ are reinitialized in the prediction step as

$$\boldsymbol{m}_{k|k-1}^{\mathrm{e},c}=\boldsymbol{m}_0^{\mathrm{e},c},\ \boldsymbol{P}_{k|k-1}^{\mathrm{e},c}=\boldsymbol{P}_0^{\mathrm{e},c}, \tag{28}$$

where $\boldsymbol{m}_0^{\mathrm{e},c}$ and $\boldsymbol{P}_0^{\mathrm{e},c}$ are the mean and covariance matrices of the prior class-dependent extent state $\boldsymbol{x}_0^{\mathrm{e},c}$, respectively.

Proposition 2 If the predicted class-dependent state is a Gaussian–Gaussian distribution as in Eq. (25), the updated state also follows a Gaussian–Gaussian distribution as

$$p(\underline{\boldsymbol{x}}_k\,|\,c,\underline{\mathcal{Z}}^k)=\mathcal{N}(\boldsymbol{x}_k^{\mathrm{e}};\boldsymbol{m}_k^{\mathrm{e},c},\boldsymbol{P}_k^{\mathrm{e},c})\mathcal{N}(\boldsymbol{x}_k^{\mathrm{k}};\boldsymbol{m}_k^{\mathrm{k}},\boldsymbol{P}_k^{\mathrm{k}}). \tag{29}$$

In our proposed JTC-RHM method, according to decomposition (20) of the posterior class-dependent density $p(\underline{\boldsymbol{x}}_k\,|\,c,\underline{\mathcal{Z}}^k)$, we first use the mean $\overline{\boldsymbol{z}}_k$ of the received measurements $\mathcal{Z}_k$ to update the predicted kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ by the Kalman-like filter, where the update kinematic state is $\mathcal{N}(\boldsymbol{x}_k^{\mathrm{k}};\boldsymbol{m}_k^{\mathrm{k}},\boldsymbol{P}_k^{\mathrm{k}})$ with

$$\boldsymbol{m}_k^{\mathrm{k}}=\boldsymbol{m}_{k|k-1}^{\mathrm{k}}+\boldsymbol{K}_{k|k-1}^{\mathrm{k}}(\overline{\boldsymbol{z}}_k-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k}}), \tag{30}$$

$$\tilde{\boldsymbol{P}}_k^{\mathrm{k}}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k}}-\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}}(\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}})^{\mathrm{T}}, \tag{31}$$

where

$$\boldsymbol{K}_{k|k-1}^{\mathrm{k}}=\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}\otimes\boldsymbol{I}_2,\ \tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k}}\tilde{\boldsymbol{H}}^{\mathrm{T}}(\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}})^{-1}, \tag{32}$$

$$\tilde{\boldsymbol{S}}_{k|k-1}=\tilde{\boldsymbol{H}}\tilde{\boldsymbol{P}}_{k|k-1}\tilde{\boldsymbol{H}}^{\mathrm{T}}+\frac{[\det(\boldsymbol{B}_k)]^{2/d}}{n_{k,z}}, \tag{33}$$

$$\overline{\boldsymbol{z}}_k=\frac{1}{n_{k,z}}\sum_{\boldsymbol{z}_k^m\in\mathcal{Z}_k}\boldsymbol{z}_k^m, \tag{34}$$

where $\tilde{\boldsymbol{H}}$ is the 1D representation of $\boldsymbol{H}$.

The prior class-dependent measurements and the received measurements are then processed sequentially to update the extent state. Since the prior class-dependent measurements change as the target moves, the class-dependent measurements $\mathcal{Z}_k^c$ is obtained by the updated kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ as

$$\begin{aligned}\mathcal{Z}_k^c&=\begin{bmatrix}\cos(\theta_k^{\mathrm{k}}) & -\sin(\theta_k^{\mathrm{k}})\\ \sin(\theta_k^{\mathrm{k}}) & \cos(\theta_k^{\mathrm{k}})\end{bmatrix}\times\mathcal{Z}^c,\\ \theta_k^{\mathrm{k}}&=\arctan\left(\frac{\boldsymbol{m}_k^{\mathrm{k}}(4)}{\boldsymbol{m}_k^{\mathrm{k}}(3)}\right).\end{aligned} \tag{35}$$

As the observation model (8) is nonlinear, neither the likelihood functions of these class-dependent measurements nor the received measurements have an explicit expression, and Eq. (12) has no explicit expression, either. Therefore, we resort to the UKF to estimate the target extent state.

We use target position component $\boldsymbol{x}_k^{\mathrm{p}}$ of the update kinematic state $\boldsymbol{x}_k^{\mathrm{k}}$ to update the extent state, and denote it as $\boldsymbol{x}_k^{\mathrm{pe}}\triangleq[(\boldsymbol{x}_k^{\mathrm{p}})^{\mathrm{T}},(\boldsymbol{x}_k^{\mathrm{e}})^{\mathrm{T}}]^{\mathrm{T}}$. For the augment vector $\boldsymbol{x}_k^{\mathrm{aug}}\triangleq[(\boldsymbol{x}_k^{\mathrm{pe}})^{\mathrm{T}},s_k,\boldsymbol{v}_k^{\mathrm{T}}]^{\mathrm{T}}$, its mean and covariance are given as

$$\boldsymbol{m}_{k|k-1}^{\mathrm{aug},c}=\begin{bmatrix}\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c}\\ \mu_s\\ \boldsymbol{0}_{2\times1}\end{bmatrix},\ \boldsymbol{P}_{k|k-1}^{\mathrm{aug},c}=\begin{bmatrix}\boldsymbol{P}_{k|k-1}^{\mathrm{pe},c} & & \\ & \sigma_s^2 & \\ & & \boldsymbol{R}_k\end{bmatrix}, \tag{36}$$

where

$$\begin{cases}\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c}=[(\boldsymbol{m}_k^{\mathrm{p}})^{\mathrm{T}},(\boldsymbol{m}_{k|k-1}^{\mathrm{e},c})^{\mathrm{T}}]^{\mathrm{T}},\\ \boldsymbol{P}_{k|k-1}^{\mathrm{pe},c}=\mathrm{blkdiag}(\boldsymbol{I}_2,\boldsymbol{P}_{k|k-1}^{\mathrm{e},c}).\end{cases} \tag{37}$$

Unscented transformation (UT) is performed on the augment vector to obtain $2n^{\mathrm{aug}}+1$ sampling points, where n^{aug} is the dimension of augment vector $\boldsymbol{x}_k^{\mathrm{aug}}$.

Use sampling point set $\mathcal{M}_{k|k-1}^{\mathrm{aug},c}=\{\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{aug},c}\}_{i=0}^{2n^{\mathrm{aug}}}$ and the corresponding weight sets to calculate the mean $\mathcal{W}_{k|k-1}^{\mathrm{m},c}=\{\tilde{\omega}_{k|k-1,i}^{\mathrm{m},c}\}_{i=0}^{2n^{\mathrm{aug}}}$ and covariance matrix $\mathcal{W}_{k|k-1}^{\mathrm{P},c}=\{\tilde{\omega}_{k|k-1,i}^{\mathrm{P},c}\}_{i=0}^{2n^{\mathrm{aug}}}$ as

$$\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{aug},c}=\begin{cases}\boldsymbol{m}_{k|k-1}^{\mathrm{aug},c}, & i=0,\\ \boldsymbol{m}_{k|k-1}^{\mathrm{aug},c}+(\sqrt{(n^{\mathrm{aug}}+\gamma)P_{k|k-1}^{\mathrm{aug},c}})_i, & i=1,2,\cdots,n^{\mathrm{aug}},\\ \boldsymbol{m}_{k|k-1}^{\mathrm{aug},c}-(\sqrt{(n^{\mathrm{aug}}+\gamma)P_{k|k-1}^{\mathrm{aug},c}})_i, & i=n^{\mathrm{aug}}+1,n^{\mathrm{aug}}+2,\cdots,2n^{\mathrm{aug}},\end{cases} \tag{38}$$

$$\tilde{\omega}_{k|k-1,i}^{\mathrm{m},c}=\begin{cases}\dfrac{\gamma}{n^{\mathrm{aug}}+\gamma}, & i=0,\\ \dfrac{1}{2(n^{\mathrm{aug}}+\gamma)}, & i=1,2,\cdots,2n^{\mathrm{aug}},\end{cases} \tag{39}$$

$$\tilde{\omega}_{k|k-1,i}^{\mathrm{P},c}=\begin{cases}\dfrac{\gamma}{(n^{\mathrm{aug}}+\gamma)}+(1-\alpha^2+\beta), & i=0,\\ \dfrac{1}{2(n^{\mathrm{aug}}+\gamma)}, & i=1,2,\cdots,2n^{\mathrm{aug}},\end{cases} \tag{40}$$

where $(\sqrt{\boldsymbol{P}})_i$ denotes the i^{th} column of the square root of matrix $\boldsymbol{P}$. Here, $\gamma=\alpha^2(n^{\mathrm{aug}}+\kappa)-n^{\mathrm{aug}}$, α, κ, and β are the parameters which control weight distribution.

According to the observation model (8), the predicted pseudo-measurement set of these sampling points is $\mathcal{Y}_{k|k-1}=\{\tilde{\boldsymbol{y}}_{k|k-1,i}\}_{i=0}^{2n^{\mathrm{aug}}}$, where

$$\tilde{\boldsymbol{y}}_{k|k-1,i}=h^{*}(\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{pe},c},\boldsymbol{v}_{k,l},s_{k,l},\boldsymbol{z}_{k,l}), \tag{41}$$

where $\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{pe},c}\triangleq\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{aug}}(1:n^{\mathrm{aug}}-3)$.

Using the real value of the pseudo-measurement 0, the mean $\boldsymbol{m}_k^{\mathrm{pe},c}$ and covariance matrix $P_k^{\mathrm{pe},c}$ of the updated state $\boldsymbol{x}_k^{\mathrm{pe}}$ are

$$\boldsymbol{m}_k^{\mathrm{pe},c}=\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c}+\boldsymbol{K}_k(0-\hat{\boldsymbol{z}}_{k|k-1}), \tag{42}$$

$$\boldsymbol{P}_k^{\mathrm{pe},c}=\boldsymbol{P}_{k|k-1}^{\mathrm{pe},c}-\boldsymbol{K}_k\boldsymbol{P}_{y_ky_k}(\boldsymbol{K}_k^{\mathrm{pe}})^{\mathrm{T}}, \tag{43}$$

where

$$\boldsymbol{K}_k=\boldsymbol{P}_{x_ky_k}\boldsymbol{P}_{y_ky_k}^{-1}, \tag{44}$$

$$\boldsymbol{P}_{y_ky_k}=\sum_{i=0}^{2n^{\mathrm{aug}}}\tilde{\omega}_{k|k-1,i}^{\mathrm{P},c}[\tilde{\boldsymbol{y}}_{k|k-1,i}-\hat{\boldsymbol{z}}_{k|k-1}][\tilde{\boldsymbol{y}}_{k|k-1,i}-\hat{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}}, \tag{45}$$

$$\boldsymbol{P}_{x_ky_k}=\sum_{i=0}^{2n^{\mathrm{aug}}}\tilde{\omega}_{k|k-1,i}^{\mathrm{P},c}[\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{pe},c}-\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c}][\tilde{\boldsymbol{y}}_{k|k-1,i}-\hat{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}}, \tag{46}$$

$$\hat{\boldsymbol{z}}_{k|k-1}=\sum_{i=0}^{2n^{\mathrm{aug}}}\tilde{\omega}_{k|k-1,i}^{\mathrm{m},c}\tilde{\boldsymbol{y}}_{k|k-1,i}. \tag{47}$$

Following Eqs. (36)–(47), the updated extent state $\boldsymbol{x}_k^{\mathrm{e}}$ can be obtained by treating these measurements sequentially. For the target class probability, the update equation of $\tilde{\mu}_k^c$ is the same as Eq. (17).

The output of state estimation at time k is calculated as

$$\begin{aligned}\hat{\underline{\boldsymbol{x}}}_k&\triangleq E[\underline{\boldsymbol{x}}_k\mid\underline{\mathcal{Z}}^k]\\&=\sum_{c=1}^{n_c}p(\underline{\boldsymbol{x}}_k\mid c,\underline{\mathcal{Z}}^k)p(c\mid\underline{\mathcal{Z}}^k)\\&=\left(\boldsymbol{m}_k^{\mathrm{k}},\sum_{i=1}^{n_c}\boldsymbol{m}_k^{\mathrm{e},c}\tilde{\mu}_k^c\right).\end{aligned} \tag{48}$$

The output of the target class probability is given as $p(c\mid\underline{\mathcal{Z}}^k)=\tilde{\mu}_k^c$.

4 The proposed JTC-RHM-Ber filter

For the tracking of a single extended target, the target state is time-varying due to the existence of target birth and death. The received measurements may be produced by clutter at each point of time. The random finite set (RFS) theoretical framework is a suitable means to solve such problems (Mahler, 2007, 2014). Based on RFS, the Bernoulli filter consists of the exact Bayes filter for a single dynamic system. For details of the Bernoulli filter, please refer to Ristic et al. (2013).

In this section, we integrate the proposed JTC-RHM method into the Bernoulli filter framework to present a JTC-RHM-Ber filter to track and classify a single extended target jointly in the presence of detection uncertainty and clutter.

For the proposed JTC-RHM-Ber filter, an extended target at time $k-1$ is modeled as $\mathcal{X}_{k-1}=\varnothing$ or $\mathcal{X}_{k-1}=\{\xi_{k-1}\}$. The existence probability is denoted as q_{k-1} and $\xi_{k-1}\triangleq(\underline{\boldsymbol{x}}_{k-1},c)$. The estimation is implemented by a Gaussian–Gaussian mixture as

$$\mathcal{S}_{k-1}(\xi)=\mathcal{S}_{k-1}(\underline{\boldsymbol{x}}\,|\,c)p(c)=\sum_{j=1}^{J_{k-1}}\begin{bmatrix} w_{k-1}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k-1}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k-1}^{\mathrm{e},c,(j)}) \\ \cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k-1}^{\mathrm{k},(j)},\boldsymbol{P}_{k-1}^{\mathrm{k},(j)})\tilde{\mu}_{k-1}^{c,(j)}\end{bmatrix}, \quad (49)$$

where J_{k-1} is the number of Gaussian–Gaussian components, $w_{k-1}^{(j)}$ is the weight of the j^{th} Gaussian–Gaussian component, and $\tilde{\mu}_{k-1}^{c,(j)}$ is the probability of the j^{th} Gaussian–Gaussian component with target class label c.

Remark 3 The probability of target existence should not be affected by the prior class-dependent measurements and target classes. The kinematic state and extent state can be estimated separately according to Eq. (20). Therefore, we use only the received measurements to calculate the existence probability and the weight of each Gaussian–Gaussian component. Covariance matrix $\boldsymbol{P}_{k-1}^{\mathrm{k},(j)}$ of kinematic state $\boldsymbol{x}^{\mathrm{k}}$ is modeled in the same way as in the JTC-RHM method; i.e., $\boldsymbol{P}_{k-1}^{\mathrm{k},(j)}=\tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k},(j)}\otimes\boldsymbol{X}_{k-1}^{(j)}$, where $\tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k},(j)}$ is the 1D representation of $\boldsymbol{P}_{k-1}^{\mathrm{k},(j)}$ and $\boldsymbol{X}_{k-1}^{(j)}$ describes the target extension. Parameter $\boldsymbol{X}_{k-1}^{(j)}$ follows the inverse Wishart distribution with two parameters $v_{k-1}^{(j)}$ and $\boldsymbol{V}_{k-1}^{(j)}$. Different from the JTC-RHM method, $\boldsymbol{X}_{k-1}^{(j)}$ is used to obtain the existence probability and the weight of each Gaussian–Gaussian component. Consequently, additional parameters $v_{k-1}^{(j)}$ and $\boldsymbol{V}_{k-1}^{(j)}$ are needed in the proposed JTC-RHM-Ber filter.

The recursive expressions for the proposed JTC-GIW-Ber filter are derived under the following assumptions:

Assumption 1 The kinematic state follows a linear Gaussian dynamical model (1) and a linear Gaussian observation model (18).

Assumption 2 The clutter is modeled as a Poisson RFS with the Poisson average rate λ_c, and is independent of target-originated measurements. The spatial distribution of clutter is uniform, denoted by $\mathcal{C}(\boldsymbol{z})$.

Assumption 3 Target-originated measurements are modeled as a Binomial RFS; they are independent of one another.

Assumption 4 Survival and detection probabilities are state-independent; i.e., $p_{\mathrm{S}}(\xi_{k-1})=p_{\mathrm{S}}$ and $p_{\mathrm{D}}(\xi_k)=p_{\mathrm{D}}$.

Assumption 5 The birth probability and the PDF of birth targets are p_{B} and a mixture of Gaussian–Gaussian components at time k, respectively:

$$S_{k,\mathrm{B}}(\xi)=S_{k,\mathrm{B}}(\underline{\boldsymbol{x}}\,|\,c)p(c)=\sum_{j=1}^{J_{k,\mathrm{B}}}\Big[w_{k,\mathrm{B}}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k,\mathrm{B}}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k,\mathrm{B}}^{\mathrm{e},c,(j)})\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k,\mathrm{B}}^{\mathrm{k},(j)},\boldsymbol{P}_{k,\mathrm{B}}^{\mathrm{k},(j)})\tilde{\mu}_{k,\mathrm{B}}^{c,(j)}\Big], \quad (50)$$

where $J_{k,\mathrm{B}}$ is the number of Gaussian–Gaussian components, $w_{k,\mathrm{B}}^{(j)}$ is the weight of the j^{th} Gaussian–Gaussian component, and $\tilde{\mu}_{k,\mathrm{B}}^{c,(j)}=1/n_c$ is the class probability of a new birth target with target class label c. Equation $\boldsymbol{P}_{k,\mathrm{B}}^{\mathrm{k},(j)}=\tilde{\boldsymbol{P}}_{k,\mathrm{B}}^{\mathrm{k},(j)}\otimes\boldsymbol{X}_{k,\mathrm{B}}^{(j)}$ and two parameters $v_{k,\mathrm{B}}^{(j)}$ and $\boldsymbol{V}_{k,\mathrm{B}}^{(j)}$ are used to describe $\boldsymbol{X}_{k,\mathrm{B}}^{(j)}$.

The proposed JTC-RHM-Ber filter includes two steps: prediction and update. In this study, the superscript $(j,\mathcal{W})$ denotes the j^{th} updated Gaussian–Gaussian component using measurement set $\mathcal{W}$.

Proposition 3 It is assumed that the spatial PDF $\mathcal{S}_{k-1}(\xi)$ is a Gaussian–Gaussian mixture at time $k-1$ (Eq. (49)). The prediction step of the proposed JTC-RHM-Ber filter is then given by

$$\mathcal{S}_{k|k-1}(\xi)=\mathcal{S}_{k|k-1}(\xi\,|\,c)p(c)=\frac{p_{\mathrm{B}}(1-q_{k-1})\mathcal{S}_{k,\mathrm{B}}(\xi)}{q_{k|k-1}}+\frac{p_{\mathrm{S}}q_{k-1}}{q_{k|k-1}}\sum_{j=1}^{J_{k-1}}\Big[w_{k-1}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k|k-1}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k|k-1}^{\mathrm{e},c,(j)})\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)},\boldsymbol{P}_{k|k-1}^{\mathrm{k},(j)})\tilde{\mu}_{k-1}^{c,(j)}\Big], \quad (51)$$

$$q_{k|k-1}=p_{\mathrm{B}}(1-q_{k-1})+p_{\mathrm{S}}q_{k-1}, \quad (52)$$

where the first term on the right side of Eq. (51) corresponds to the prediction of birth targets, denoted as $\mathcal{S}_{k|k-1,\mathrm{B}}(\xi)$, while the second term is the prediction of the existing targets, denoted as $\mathcal{S}_{k|k-1,\mathrm{S}}(\xi)$. $\mathcal{S}_{k|k-1,\mathrm{B}}(\xi)$ and $\mathcal{S}_{k|k-1,\mathrm{S}}(\xi)$ can be calculated as

$$\mathcal{S}_{k|k-1,\mathrm{B}}(\xi)=\sum_{j=1}^{J_{k,\mathrm{B}}}\Big[w_{k|k-1,\mathrm{B}}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k|k-1,\mathrm{B}}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k|k-1,\mathrm{B}}^{\mathrm{e},c,(j)})\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k|k-1,\mathrm{B}}^{\mathrm{k},(j)},\boldsymbol{P}_{k|k-1,\mathrm{B}}^{\mathrm{k},(j)})\tilde{\mu}_{k,\mathrm{B}}^{c,(j)}\Big], \quad (53)$$

$$\mathcal{S}_{k|k-1,\mathrm{S}}(\xi)=\sum_{j=1}^{J_{k-1}}\Big[w_{k|k-1,\mathrm{S}}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k|k-1,\mathrm{S}}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k|k-1,\mathrm{S}}^{\mathrm{e},c,(j)})\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k|k-1,\mathrm{S}}^{\mathrm{k},(j)},\boldsymbol{P}_{k|k-1,\mathrm{S}}^{\mathrm{k},(j)})\tilde{\mu}_{k,\mathrm{S}}^{c,(j)}\Big]. \tag{54}$$

For the prediction of new birth targets, the parameters of $\mathcal{S}_{k|k-1,\mathrm{B}}(\xi)$ are given by

$$w_{k|k-1,\mathrm{B}}^{(j)}=\frac{p_{\mathrm{B}}(1-q_{k-1})}{q_{k|k-1}}w_{k,\mathrm{B}}^{(j)}, \tag{55}$$

$$\boldsymbol{m}_{k|k-1,\mathrm{B}}^{\mathrm{k},(j)}=\boldsymbol{m}_{k,\mathrm{B}}^{\mathrm{k},(j)},\ \tilde{\boldsymbol{P}}_{k|k-1,\mathrm{B}}^{\mathrm{k},(j)}=\tilde{\boldsymbol{P}}_{k,\mathrm{B}}^{\mathrm{k},(j)}, \tag{56}$$

$$v_{k|k-1,\mathrm{B}}^{(j)}=v_{k,\mathrm{B}}^{(j)},\ \boldsymbol{V}_{k|k-1,\mathrm{B}}^{(j)}=\boldsymbol{V}_{k,\mathrm{B}}^{(j)}, \tag{57}$$

$$\boldsymbol{m}_{k|k-1,\mathrm{B}}^{\mathrm{e},c,(j)}=\boldsymbol{m}_{k,\mathrm{B}}^{\mathrm{e},c,(j)},\ \boldsymbol{P}_{k|k-1,\mathrm{B}}^{\mathrm{e},c,(j)}=\boldsymbol{P}_{k,\mathrm{B}}^{\mathrm{e},c,(j)}. \tag{58}$$

For the prediction of existing targets, the process is similar to that of the JTC-RHM method. For each Gaussian–Gaussian component, the kinematic parameters of $\mathcal{S}_{k|k-1,\mathrm{S}}(\xi)$ are given by

$$w_{k|k-1,\mathrm{S}}^{(j)}=\frac{p_{\mathrm{S}}q_{k-1}}{q_{k|k-1}}w_{k|k-1}^{(j)}, \tag{59}$$

$$\boldsymbol{m}_{k|k-1,\mathrm{S}}^{\mathrm{k},(j)}=F_k^{\mathrm{k}}\boldsymbol{m}_{k-1}^{\mathrm{k},(j)}, \tag{60}$$

$$\tilde{\boldsymbol{P}}_{k|k-1,\mathrm{S}}^{\mathrm{k},(j)}=\tilde{\boldsymbol{F}}_k^{\mathrm{k}}\tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k},(j)}(\tilde{\boldsymbol{F}}_k^{\mathrm{k}})^{\mathrm{T}}+\tilde{\boldsymbol{Q}}_k^{\mathrm{k}}, \tag{61}$$

$$v_{k|k-1}^{(j)}=\frac{2\delta_k(\lambda_{k-1}^{(j)}+1)(\lambda_{k-1}^{(j)}-1)(\lambda_{k-1}^{(j)}-2)}{(\lambda_{k-1}^{(j)})^2(\lambda_{k-1}^{(j)}+\delta_k)}+2d+4, \tag{62}$$

$$\boldsymbol{V}_{k|k-1}^{(j)}=\frac{\delta_k(v_{k|k-1,\mathrm{S}}^{(j)}-2d-2)}{\lambda_{k-1}^{(j)}}\boldsymbol{A}_k\boldsymbol{V}_{k-1}^{(j)}\boldsymbol{A}_k^{\mathrm{T}}, \tag{63}$$

where $\lambda_{k-1}^{(j)}=v_{k-1}^{(j)}-2d-2$, and δ_k is a scalar. The number of Gaussian–Gaussian mixtures for the prediction target state is $J_{k|k-1}=J_{k-1}+J_{k,\mathrm{B}}$.

For prediction of the extent state, the mean and covariance matrices of the extent state are reinitialized as

$$\begin{cases}\boldsymbol{m}_{k|k-1,\mathrm{S}}^{\mathrm{e},c,(j)}=\boldsymbol{m}_0^{\mathrm{e},c},\\ \boldsymbol{P}_{k|k-1,\mathrm{S}}^{\mathrm{e},c,(j)}=\boldsymbol{P}_0^{\mathrm{e},c}.\end{cases} \tag{64}$$

The Bernoulli filter for extended targets models target-originated measurements using binomial RFS, while the cardinality of this RFS follows a binomial distribution with parameter l. The number of target-originated measurements is denoted as l_k at time k. Due to detection uncertainty, zero, one, or multiple target-originated measurements can be detected, meaning that the cardinality of received measurements is $|\mathcal{W}|=0, 1, \ldots, l_k$. We use distance partition (Granström et al., 2012) to obtain all subsets of the received target-originated measurements, and denote subsets as $\boldsymbol{P}_{1:lk}(\mathcal{Z}_k)$. The method in Eryildirim and Guldogan (2016) is used to obtain l_k.

Proposition 4 If the predicted spatial PDF $\mathcal{S}_{k|k-1}(\xi)$ of the proposed JTC-RHM-Ber filter is a Gaussian–Gaussian mixture:

$$\begin{aligned}\mathcal{S}_{k|k-1}(\xi)&=\mathcal{S}_{k|k-1}(\xi\,|\,c)p(c)\\&=\sum_{j=1}^{J_{k|k-1}}\Big[w_{k|k-1}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k|k-1}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k|k-1}^{\mathrm{e},c,(j)})\\&\quad\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)},\boldsymbol{P}_{k|k-1}^{\mathrm{k},(j)})\tilde{\mu}_{k-1}^{c,(j)}\Big],\end{aligned} \tag{65}$$

then the updated spatial PDF $\mathcal{S}_k(\xi)$ is also a Gaussian–Gaussian mixture. The updated equations are given by

$$\begin{aligned}\mathcal{S}_k(\xi)&=\frac{(1-p_{\mathrm{D}})^{l_k}\mathcal{S}_{k|k-1}(\xi)}{1-\Delta_k}\\&\quad+\frac{\sum\limits_{\mathcal{W}\in P_{1:l_k}(\mathcal{Z}_k)}\psi_k\prod\limits_{z\in\mathcal{W}}\frac{g_k(z\,|\,\xi)}{\lambda_{\mathrm{c}}\mathcal{C}(z)}}{1-\Delta_k}\mathcal{S}_{k|k-1}(\xi),\end{aligned} \tag{66}$$

$$q_k=\frac{1-\Delta_k}{1-q_{k|k-1}\Delta_k}q_{k|k-1}, \tag{67}$$

with

$$\Delta_k=1-(1-p_{\mathrm{D}})^{l_k}-\sum_{\mathcal{W}\in P_{1:l_k}(\mathcal{Z}_k)}\psi_k\frac{\int\prod\limits_{z\in\mathcal{W}}g_k(z\,|\,\xi)\mathcal{S}_{k|k-1}(\xi)\mathrm{d}\xi}{\prod\limits_{z\in\mathcal{W}}\lambda_{\mathrm{c}}\mathcal{C}(z)}, \tag{68}$$

$$\psi_k=\frac{l_k!}{(l_k-|\mathcal{W}|)!}\frac{p_{\mathrm{D}}^{|\mathcal{W}|}}{(1-p_{\mathrm{D}})^{|\mathcal{W}|-l_k}}, \tag{69}$$

where $g_k(z|\xi)$ is the likelihood function, which can be obtained via the observation model (18).

According to Eq. (66), the posterior spatial PDF $\mathcal{S}_k(\xi)$ consists of two parts, the non-detection case $\mathcal{S}_{k,\mathrm{ND}}(\xi)$ and the detection case $\mathcal{S}_{k,\mathrm{D}}(\xi,\mathcal{W})$, namely

$$\mathcal{S}_k(\xi)=\mathcal{S}_{k,\mathrm{ND}}(\xi)+\sum_{\mathcal{W}\in P_{1:J_k}(\mathcal{Z}_k)}\mathcal{S}_{k,\mathrm{D}}(\xi,\mathcal{W}). \tag{70}$$

For the posterior spatial PDF of the non-detection case, $\mathcal{S}_{k,\mathrm{ND}}(\xi)$ is a Gaussian–Gaussian mixture, obtained by

$$\begin{aligned}S_{k,\mathrm{ND}}(\xi)&=\frac{(1-p_{\mathrm{D}})^{l_k}\mathcal{S}_{k|k-1}(\xi)}{1-\Delta_k}\\&=\sum_{j=1}^{J_{k|k-1}}\Big[w_{k,\mathrm{ND}}^{(j)}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k,\mathrm{ND}}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k,\mathrm{ND}}^{\mathrm{e},c,(j)})\\&\quad\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k,\mathrm{ND}}^{\mathrm{k},(j)},\boldsymbol{P}_{k,\mathrm{ND}}^{\mathrm{k},(j)})\tilde{\mu}_{k-1}^{c,(j)}\Big],\end{aligned} \tag{71}$$

with

$$w_{k,\mathrm{ND}}^{(j)}=\frac{(1-p_{\mathrm{D}})^{l_k}}{1-\Delta_k}w_{k|k-1}^{(j)}, \tag{72}$$

$$\boldsymbol{m}_{k,\mathrm{ND}}^{\mathrm{k},(j)}=\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)},\tilde{\boldsymbol{P}}_{k,\mathrm{ND}}^{\mathrm{k},(j)}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(j)}, \tag{73}$$

$$v_{k,\mathrm{ND}}^{(j)}=v_{k|k-1}^{(j)},\boldsymbol{V}_{k,\mathrm{ND}}^{(j)}=\boldsymbol{V}_{k|k-1}^{(j)}, \tag{74}$$

$$\boldsymbol{m}_{k,\mathrm{ND}}^{\mathrm{e},c,(j)}=\boldsymbol{m}_{k|k-1}^{\mathrm{e},c,(j)},\boldsymbol{P}_{k,\mathrm{ND}}^{\mathrm{e},c,(j)}=\boldsymbol{P}_{k|k-1}^{\mathrm{e},c,(j)}. \tag{75}$$

For the posterior spatial PDF of the detection case, $\mathcal{S}_{k,\mathrm{D}}(\xi,\mathcal{W})$ is also a Gaussian–Gaussian mixture, obtained by

$$\begin{aligned}\mathcal{S}_{k,\mathrm{D}}(\xi,\mathcal{W})&=\frac{\psi_k\prod_{z\in\mathcal{W}}\frac{g_k(z\,|\,\xi)}{\lambda_{\mathrm{c}}\mathcal{C}(z)}}{1-\Delta_k}\mathcal{S}_{k|k-1}(\xi)\\&=\sum_{j=1}^{J_{k|k-1}}\Big[w_{k,\mathrm{D}}^{(j,\mathcal{W})}\mathcal{N}(\boldsymbol{x}^{\mathrm{e}};\boldsymbol{m}_{k,\mathrm{D}}^{\mathrm{e},c,(j,\mathcal{W})},\boldsymbol{P}_{k,\mathrm{D}}^{\mathrm{e},c,(j,\mathcal{W})})\\&\quad\cdot\mathcal{N}(\boldsymbol{x}^{\mathrm{k}};\boldsymbol{m}_{k,\mathrm{D}}^{\mathrm{k},(j,\mathcal{W})},\boldsymbol{P}_{k,\mathrm{D}}^{\mathrm{k},(j,\mathcal{W})})\tilde{\mu}_{k-1,\mathrm{D}}^{c,(j,\mathcal{W})}\Big].\end{aligned} \tag{76}$$

Firstly, the received measurements are used to obtain the weight and parameters of each Gaussian–Gaussian component, as well as the probability of target existence using a Kalman-like filter:

$$w_{k,\mathrm{D}}^{(j,\mathcal{W})}=\frac{1}{1-\Delta_k}\psi_k\left(\mathcal{C}(z)\lambda_{\mathrm{c}}\right)^{-|\mathcal{W}|}\mathcal{L}_k^{(j,\mathcal{W})}w_{k|k-1}^{(j)}, \tag{77}$$

$$\boldsymbol{m}_{k,\mathrm{D}}^{\mathrm{k},(j,\mathcal{W})}=\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}+\boldsymbol{K}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}(\overline{\boldsymbol{z}}_k^{(\mathcal{W})}-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}), \tag{78}$$

$$\tilde{\boldsymbol{P}}_{k,\mathrm{D}}^{\mathrm{k},(j,\mathcal{W})}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(j)}-\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}(\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})})^{\mathrm{T}}, \tag{79}$$

$$v_{k,\mathrm{D}}^{(j,\mathcal{W})}=|\,\mathcal{W}\,|+v_{k|k-1}^{(j)}, \tag{80}$$

$$\boldsymbol{V}_{k,\mathrm{D}}^{(j,\mathcal{W})}=\boldsymbol{V}_{k|k-1}^{(j)}+(\boldsymbol{B}_k^{(j)})^{-1}\overline{\boldsymbol{Z}}_k^{(\mathcal{W})}(\boldsymbol{B}_k^{(j)})^{-\mathrm{T}}+\boldsymbol{N}_k^{(j,\mathcal{W})}, \tag{81}$$

with

$$\boldsymbol{K}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}=\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}\otimes\boldsymbol{I}_2,\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(j)}\tilde{\boldsymbol{H}}^{\mathrm{T}}(\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})})^{-1}, \tag{82}$$

$$\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})}=\tilde{\boldsymbol{H}}\tilde{\boldsymbol{P}}_{k|k-1}^{(j)}\tilde{\boldsymbol{H}}^{\mathrm{T}}+\frac{[\det(\boldsymbol{B}_k)]^{2/d}}{|\,\mathcal{W}\,|}, \tag{83}$$

$$\boldsymbol{N}_k^{(j,\mathcal{W})}=(\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k},(j,\mathcal{W})})^{-1}(\overline{\boldsymbol{z}}_k^{(\mathcal{W})}-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)})(\overline{\boldsymbol{z}}_k^{(\mathcal{W})}-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)})^{\mathrm{T}}, \tag{84}$$

$$\overline{\boldsymbol{z}}_k^{(\mathcal{W})}=\frac{1}{|\,\mathcal{W}\,|}\sum_{\boldsymbol{z}_k\in\mathcal{W}}\boldsymbol{z}_k,\overline{\boldsymbol{Z}}_k^{(\mathcal{W})}=\sum_{\boldsymbol{z}_k\in\boldsymbol{W}}(\boldsymbol{z}_k-\overline{\boldsymbol{z}}_k^{(\mathcal{W})})(\boldsymbol{z}_k-\overline{\boldsymbol{z}}_k^{(\mathcal{W})})^{\mathrm{T}}, \tag{85}$$

$$\begin{aligned}\Delta_k&=1-(1-p_{\mathrm{D}})^{l_k}\\&\quad-\sum_{\mathcal{W}\in P_{1:J_k}(\mathcal{Z}_k)}\psi_k\left(\mathcal{C}(z)\lambda_{\mathrm{c}}\right)^{-|\mathcal{W}|}\sum_{j=1}^{J_{k|k-1}}\left(\mathcal{L}_k^{(j,\mathcal{W})}w_{k|k-1}^{(j,\mathcal{W})}\right),\end{aligned} \tag{86}$$

$$\begin{aligned}\mathcal{L}_k^{(j,\mathcal{W})}&=\left[\det(\boldsymbol{B}_k^{(j)})\right]^{\frac{-(|\mathcal{W}|-1)}{2}}\left((\pi)^{-|\mathcal{W}|}\,|\,\mathcal{W}\,|\det(\tilde{\boldsymbol{S}}_{k|k-1}^{(j,\mathcal{W})})\right)^{-\frac{d}{2}}\\&\quad\cdot\frac{\left[\det(\boldsymbol{V}_{k|k-1}^{(j)})\right]^{\frac{v_{k|k-1}^{(j)}-d-1}{2}}\Gamma_d\left(\frac{v_k^{(j,\mathcal{W})}-d-1}{2}\right)}{\left[\det(\boldsymbol{V}_k^{(j,\mathcal{W})})\right]^{\frac{v_k^{(j,\mathcal{W})}-d-1}{2}}\Gamma_d\left(\frac{v_{k|k-1}^{(j)}-d-1}{2}\right)},\end{aligned} \tag{87}$$

where $|\cdot|$ denotes the number of elements in a set, and $\Gamma_d(\cdot)$ is a multivariate gamma function with dimensionality d.

Subsequently, class-dependent measurements and received measurements are used to update the extent state $\boldsymbol{x}_k^{\mathrm{e}}$ using UKF, whereas parameters $\boldsymbol{m}_{k,\mathrm{D}}^{\mathrm{e},c,(j,\mathcal{W})}$ and $\boldsymbol{P}_{k,\mathrm{D}}^{\mathrm{e},c,(j,\mathcal{W})}$ can be obtained by Eqs. (36)–(47).

The update of the target class probability is conducted as

$$\tilde{\mu}_{k,\mathrm{D}}^{c,(j,\mathcal{W})}=\frac{\tilde{\eta}_{k,\mathrm{D}}^{c,(j,\mathcal{W})}\tilde{\mu}_{k|k-1}^{c,(j)}}{\sum_{c=1}^{n_{\mathrm{c}}}\tilde{\eta}_{k,\mathrm{D}}^{c,(j,\mathcal{W})}\tilde{\mu}_{k|k-1}^{c,(j)}}, \tag{88}$$

where $\tilde{\eta}_{k,\mathrm{D}}^{c,(j,\mathcal{W})}$ is the Euclidean distance of the normalized Fourier descriptors between $\boldsymbol{m}_0^{\mathrm{e},c}$ and $\boldsymbol{m}_{k,\mathrm{D}}^{\mathrm{e},c,(j,\mathcal{W})}$, which can be obtained by Eq. (16).

For the state extraction of the JTC-GIW-Ber filter, it is judged at first whether the target exists or not. The existence of the target is confirmed if $q_k\geq0.5$, while the estimated state and class probability are obtained using the maximum a posterior (MAP)

criterion with $\hat{\underline{\boldsymbol{m}}}_k=(\boldsymbol{m}_k^{\mathrm{k},(j^*)},\sum_{c=1}^{n_c}\boldsymbol{m}_k^{\mathrm{e},(j^*)}\tilde{\mu}_k^{c,(j^*)})$, $p(c)=\tilde{\mu}_k^{c,(j^*)}$, and $j^*=\arg\max_j w_k^{(j)}$.

The number of Gaussian–Gaussian components in the proposed JTC-GIW-Ber filter will grow exponentially. Therefore, the pruning and merging strategy in Granström et al. (2012) is adopted to reduce the computation burden. The main steps of the proposed JTC-GIW-Ber filter are given in the appendix.

5 Simulations

The proposed JTC-RHM method and the JTC-RHM-Ber filter are simulated to validate their effectiveness in this section. Three simulation scenarios are considered to test the performance of the proposed JTC-RHM method and the JTC-RHM-Ber filter. To illustrate the advantages of the proposed JTC-RHM method in classifying targets with complex shapes, we compare the proposed JTC-RHM method with the JTC-RMM method (Lan and Li, 2013) in simulation scenario 1. We compare the proposed JTC-RHM method with the star-convex RHM based ETT (abbreviated as RHM) method (Baum and Hanebeck, 2014) in simulation scenario 2 to illustrate the advantages of the proposed JTC-RHM in estimating the target state. Simulation scenario 3 is conducted to verify the effectiveness of the proposed JTC-RHM-Ber filter in the presence of clutter and detection uncertainty. Finally, we compare the proposed JTC-RHM method with the RHM method to analyze the algorithm complexity and time consumption.

In simulation scenarios 1 and 2, there is only one extended target in the surveillance area. In simulation scenario 3, at most one extended target appears in the surveillance area.

5.1 Simulation setup

In the three simulation scenarios, we assume that two kinds of extended targets may exist in the surveillance area. Target class 1 is composed of two orthogonal rectangles, with the length and width of the two rectangles being (34 m, 5 m) and (20 m, 5 m), respectively. The shape of target class 2 is a five-pointed star, and its circumscribed circle radius is 10 m. Prior class-dependent measurements $\{\mathcal{Z}^c\}_{c=1}^{n_c}$ and the corresponding extent state $\boldsymbol{x}_0^{\mathrm{e},c}$ are shown in Fig. 2. The class labels of targets A and B in the simulations are labels 1 and 2, respectively.

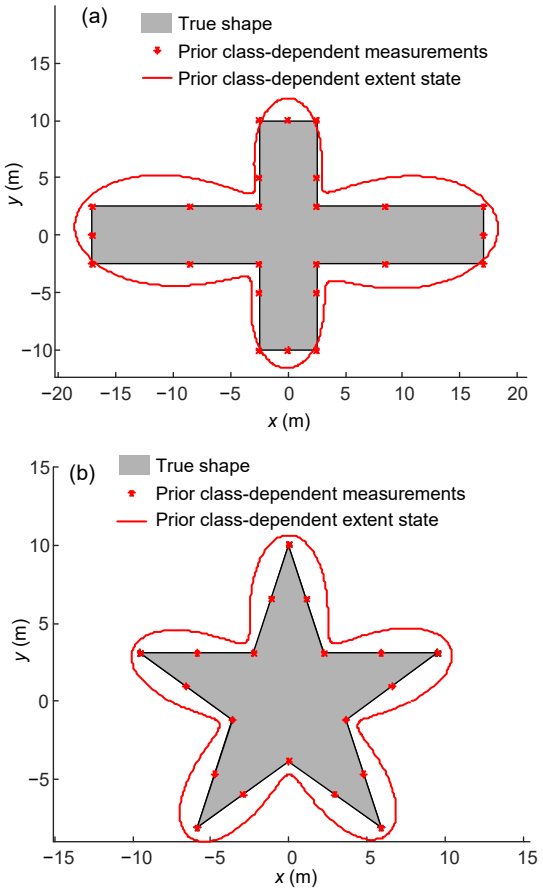

Fig. 2 The true shapes and prior information of target class 1 (a) and target class 2 (b)

In the three simulation scenarios, the trajectory of the extended target centroid is shown in Fig. 3. The blue points denote the centroid position at each sampling time point, whereas the arrow indicates the direction of target motion. The sampling interval is t=5 s. The number of received measurements follows a Poisson distribution with a measurement rate of λ_{m}=15. In scenario 3, the Poisson average clutter rate λ_{c} is set as 5, while the spatial distribution of clutter is $\mathcal{C}(z)=1/(1500\times1200)$. The covariance matrix of observation noise in the three simulation scenarios is $\boldsymbol{R}_k=\mathrm{diag}([0.1,0.1])\,\mathrm{m}^2$. For the proposed JTC-RHM method, JTC-RHM-Ber filter, and JTC-RMM method, the initial target class probabilities are $c_0^1=c_0^2=1/2$.

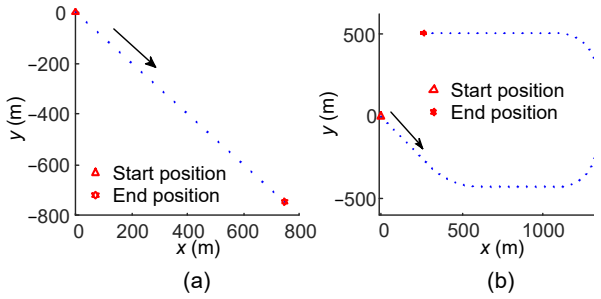

Fig. 3 The trajectory of the extended target centroid in simulation scenario 1 (a) and simulation scenarios 2 and 3 (b) (References to color refer to the online version of this figure)

The main parameters used in the proposed JTC-RHM method and the RHM method are as follows:

The order of the Fourier series is n_F=5. The process noise covariance of kinematic state is $\tilde{\boldsymbol{Q}}_k^{\mathrm{k}} = 0.5^2\,\mathrm{diag}([2,1])$. The scale factor $s_{k,l}$ follows a Gaussian distribution with a mean of 0.7 and a variance of 0.08. Initial state $\underline{\boldsymbol{x}}_0$ and its covariance $\boldsymbol{P}_0$ are $\boldsymbol{x}_0=(\boldsymbol{x}_0^{\mathrm{k}},\boldsymbol{x}_0^{\mathrm{e}})$ and $\boldsymbol{P}_0=(\boldsymbol{P}_0^{\mathrm{k}},\boldsymbol{P}_0^{\mathrm{e}})$, respectively, where $\boldsymbol{x}_0^{\mathrm{e}}=[18,\boldsymbol{0}_{10\times1}]^{\mathrm{T}}$, $\boldsymbol{x}_0^{\mathrm{k}}=[1\ \mathrm{m},\ 1\ \mathrm{m},\ 8\ \mathrm{m/s},-8\ \mathrm{m/s}]^{\mathrm{T}}$, $\boldsymbol{P}_0^{\mathrm{k}}=\mathrm{diag}([1\ \mathrm{m}^2,1\ \mathrm{m}^2,1\ \mathrm{m}^2/\mathrm{s}^2,1\ \mathrm{m}^2/\mathrm{s}^2])$, and $\boldsymbol{P}_0^{\mathrm{e}}=\mathrm{diag}([0.5\times\boldsymbol{1}_{1\times11}])$. Vector $\boldsymbol{0}_{10\times1}$ is an all-zero vector with a dimension of 10×1, while $\boldsymbol{1}_{1\times11}$ is an all-one vector with a dimension of 1×11. For the proposed JTC-RHM-Ber filter, the survival and detection probabilities are $p_{\mathrm{S}}(\xi_{k-1})$=0.99 and $p_{\mathrm{D}}(\xi_k)$=0.98, respectively. The birth model is adopted as

$$\boldsymbol{m}_{k-1,\mathrm{B}}^{\mathrm{k}}=[0\ \mathrm{m},0\ \mathrm{m},7.8\ \mathrm{m/s},-7.8\ \mathrm{m/s}]^{\mathrm{T}}, \tag{89}$$

$$\tilde{\boldsymbol{P}}_{k-1,\mathrm{B}}^{\mathrm{k}}=\mathrm{diag}([10^2\ \mathrm{m}^2,1\ \mathrm{m}^2/\mathrm{s}^2]), \tag{90}$$

$$\boldsymbol{m}_{k-1,\mathrm{B}}^{\mathrm{e},1}=[21,\boldsymbol{0}_{1\times10}]^{\mathrm{T}},\ \boldsymbol{P}_{k-1,\mathrm{B}}^{\mathrm{e},1}=2\times\mathrm{diag}(\boldsymbol{1}_{1\times11}), \tag{91}$$

$$\boldsymbol{m}_{k-1,\mathrm{B}}^{\mathrm{e},2}=[15,\boldsymbol{0}_{1\times10}]^{\mathrm{T}},\ \boldsymbol{P}_{k-1,\mathrm{B}}^{\mathrm{e},2}=2\times\mathrm{diag}(\boldsymbol{1}_{1\times11}), \tag{92}$$

$$\tilde{\mu}_{k-1,\mathrm{B}}^{1}=1/2,\ \tilde{\mu}_{k-1,\mathrm{B}}^{2}=1/2. \tag{93}$$

Further parameters of the proposed JTC-RHM-Ber filter are the same as those of the proposed JTC-RHM method.

The JTC-RMM method and the RMM method are used as a comparison group in simulation scenario 1. For the JTC-RMM method and the RMM method, the extended state is modeled by the SPD matrix $\boldsymbol{X}_k$. The target state is denoted as $(\boldsymbol{x}_{k,\mathrm{RMM}},\boldsymbol{X}_k)$ with the same parameters used by Lan and Li (2013). Therefore, the process noise covariance is $\boldsymbol{Q}_{k,\mathrm{RMM}}=(\Sigma^2(1-\mathrm{e}^{-2t/\theta})\times\mathrm{diag}([0,0,1]))\otimes\boldsymbol{X}_k$, where $\Sigma=10^{-3}\ \mathrm{m}^2/\mathrm{s}$ and $\theta=8t$. The initial kinematic state $\boldsymbol{x}_{0,\mathrm{RMM}}$ and its covariance matrix $\boldsymbol{P}_{0,\mathrm{RMM}}$ are $\boldsymbol{x}_{0,\ \mathrm{RMM}}=[1\ \mathrm{m},\ 1\ \mathrm{m},\ 8\ \mathrm{m/s},\ -8\ \mathrm{m/s},\ 0,\ 0]^{\mathrm{T}}$, and $\tilde{\boldsymbol{P}}_{0,\mathrm{RMM}}=\mathrm{diag}([10\ \mathrm{m}^2,\ 1\ \mathrm{m}^2/\mathrm{s}^2,\ 1\ \mathrm{m}^2/\mathrm{s}^4])$, respectively. The initial parameters of the extent state $\boldsymbol{X}_0$ are v_0=7 and $\boldsymbol{V}_0=\mathrm{diag}([5^2,\ 5^2])$. The adopted prior information is the class-dependent SPD matrix with $\boldsymbol{Z}^{\mathrm{p},1}=\begin{bmatrix}17^2 & 0\\ 0 & 10^2\end{bmatrix}\mathrm{m}^2$ and $\boldsymbol{Z}^{\mathrm{p},2}=\begin{bmatrix}10^2 & 0\\ 0 & 10^2\end{bmatrix}\mathrm{m}^2$.

5.2 Simulation results

We test the performance of the proposed JTC-RHM method and the JTC-RHM-Ber filter both for tracking and classification.

5.2.1 Results for simulation scenario 1

For simulation scenario 1, Figs. 4 and 5 show the tracking results of the JTC-RMM method and the proposed JTC-RHM method, respectively.

The JTC-RMM method and the RMM method have nearly the same performance in estimating the target state (kinematic and extent states) (Fig. 4). However, the proposed JTC-RHM method performs better than the RHM method in estimating the target extent state (Fig. 5); this means that the latter is more effective in describing the target shape for a target with a complex shape.

Classification results are shown in Fig. 6, where class probabilities are calculated by 100 Monte Carlo trials. As shown in Fig. 6a, the JTC-RMM method cannot correctly classify the target, where the true probability of target A should be 1; however, the estimated class probability is approximately equal to 0.1. Conversely, the proposed JTC-RHM method can obtain correct classification results (Fig. 6), where the estimated class probabilities of targets A and B are very close to the ground truth. In this simulation scenario, targets A and B have similar size information. The JTC-RMM method models a target as an elliptical shape, and cannot correctly classify the

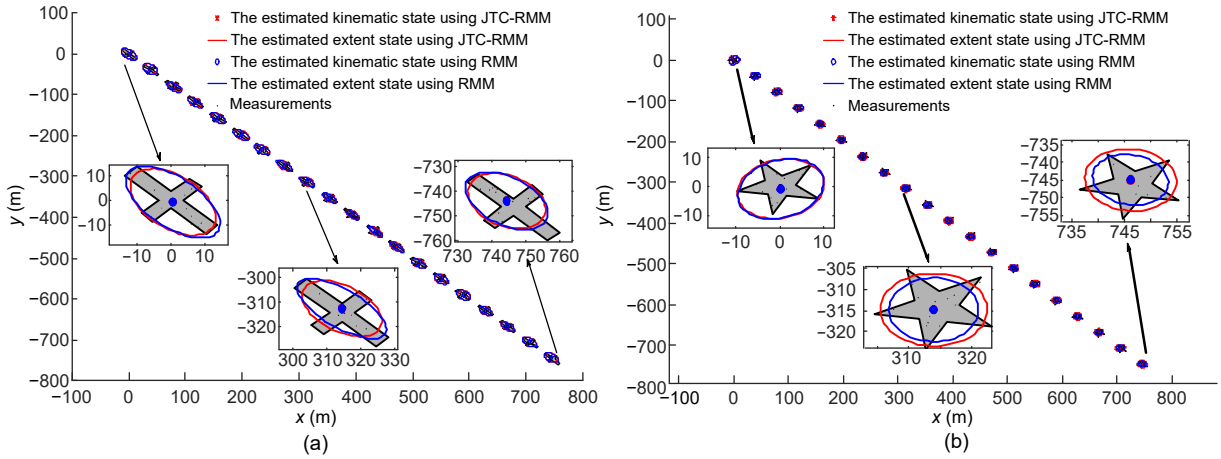

Fig. 4 Simulation results of target tracking using the JTC-RMM method for targets A (a) and B (b) in scenario 1
JTC: joint tracking and classification; RMM: random matrix model. References to color refer to the online version of this figure

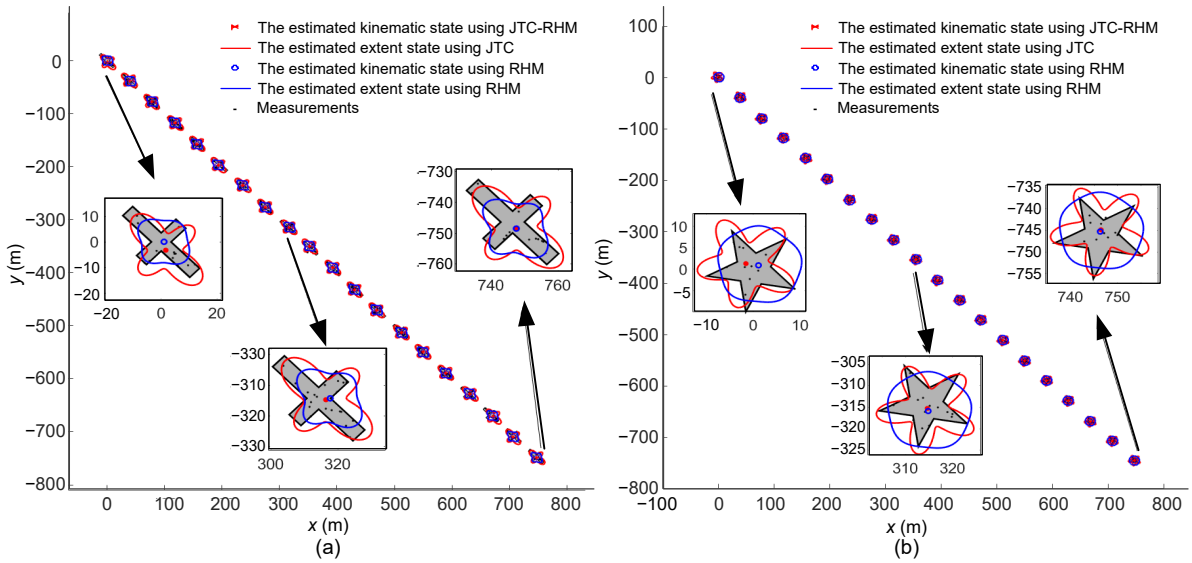

Fig. 5 Simulation results of target tracking using the proposed JTC-RHM method for targets A (a) and B (b) in scenario 1
JTC: joint tracking and classification; RHM: random hypersurface model. References to color refer to the online version of this figure

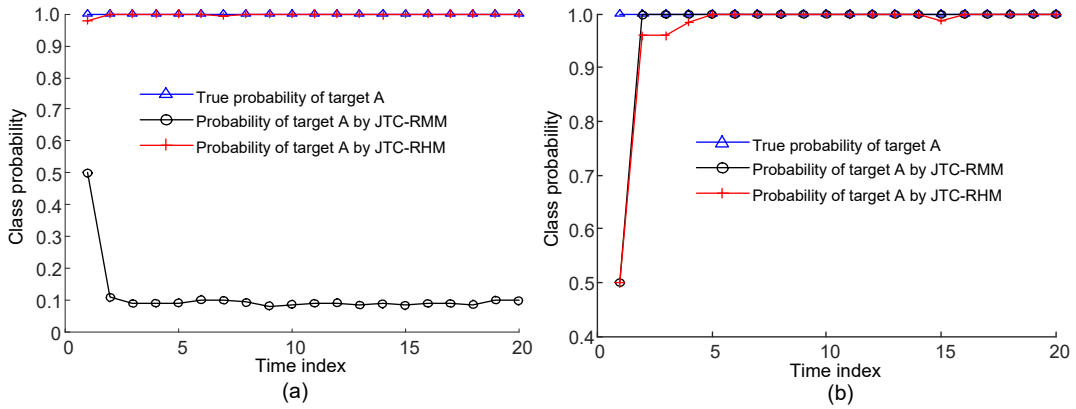

Fig. 6 Simulation results of target classification for targets A (a) and B (b) in scenario 1
JTC: joint tracking and classification; RMM: random matrix model; RHM: random hypersurface model. References to color refer to the online version of this figure

target when only the size information is used as a classification feature. In contrast, the proposed JTC-RHM method models a target as a star-convex shape and uses the contour shape as the classification feature. We can see that even if targets have similar size information, they can still be correctly classified using the proposed JTC-RHM method, as long as shapes are different.

5.2.2 Results for simulation scenario 2

Firstly, we qualitatively analyze the performance of the proposed JTC-RHM method in target tracking in scenario 2, with results shown in Fig. 7.

As shown in Fig. 7, there are eight partially enlarged subfigures, denoted by A–H. The proposed JTC-RHM method can significantly improve the estimation of the kinematic state and extent state, compared to the RHM method.

Regardless of what the true target shape is, the RHM method assumes the initial target shape as a circle, as shown in Fig. 7 marked with a blue line (denoted by subfigure A). Although the estimated shape can gradually approach the true one with the recursions of the filter, the results obtained are not very desirable. The proposed JTC-RHM method can quickly approach the true target extent state, as shown in subfigures A and B in Fig. 7, where the shape represented by the red line is more consistent with the true target shape.

As opposed to the RHM method, the proposed JTC-RHM method is more accurate in estimating the kinematic state and extent state when the target performs the maneuver (Fig. 7), where the shape represented by the red line is more consistent with the true target shape, with the red points located almost near the target centroid (such as in subfigures C–H).

Furthermore, we quantitatively evaluate the performance of the proposed JTC-RHM method on target tracking and target classification. Specifically, we use the root mean square error (RMSE) of the centroid position of the target to evaluate the kinematic state estimation. We also use the Euclidean distance of the normalized Fourier descriptors between the estimated extent state and the prior extent state (termed "shape similarity" in this study) to evaluate the extent state estimation. Quantitative evaluation results and class probabilities are obtained through 100 Monte Carlo trials, with results shown in Figs. 8 and 9. These results are shown in three aspects: kinematic state estimation, extent state estimation, and class probability.

The performance of the proposed JTC-RHM method is better than that of the RHM method for kinematic state estimation (Figs. 8a and 9a). The

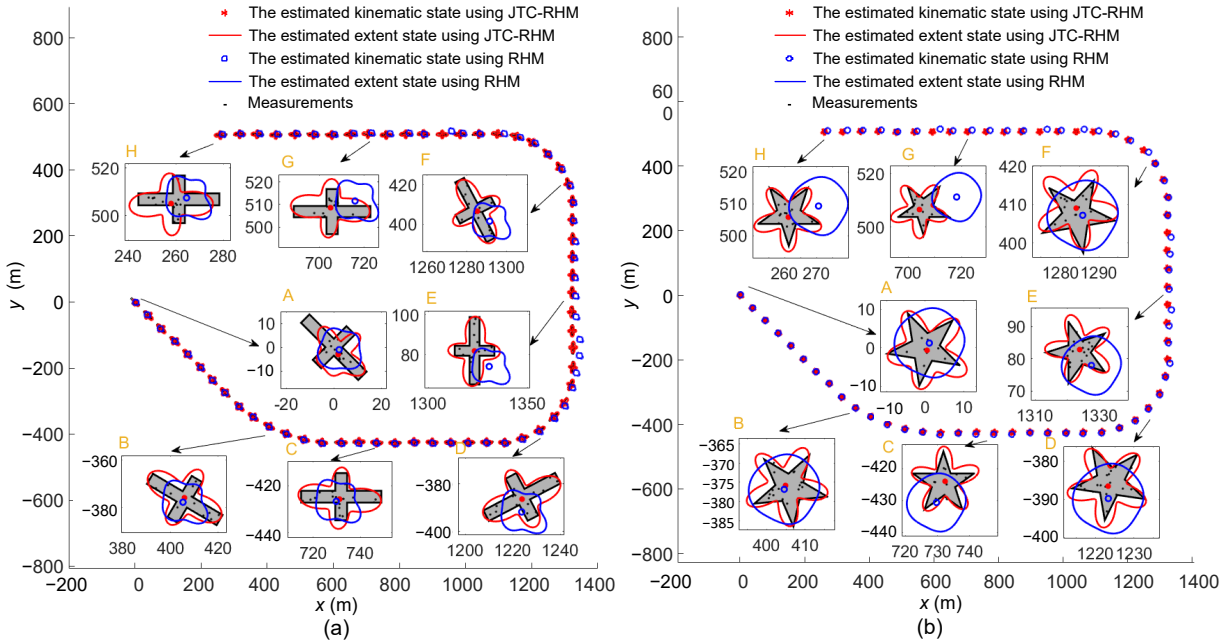

Fig. 7 Simulation results of the JTC-RHM method for targets A (a) and B (b) in scenario 2

JTC: joint tracking and classification; RHM: random hypersurface model. References to color refer to the online version of this figure

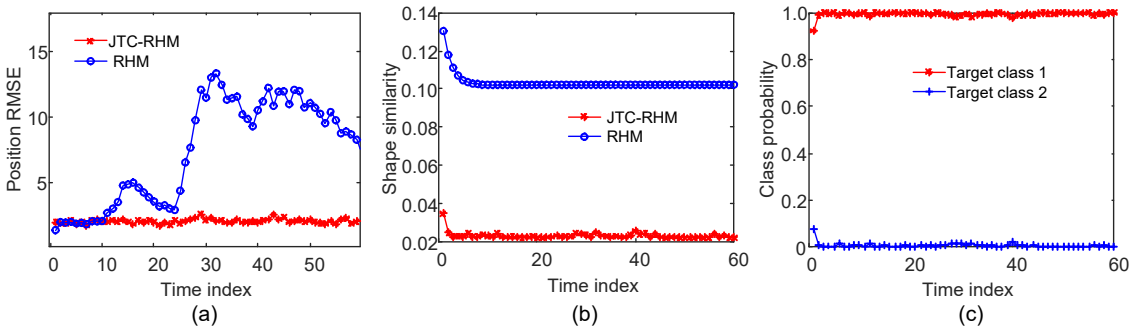

Fig. 8 Simulation results for target A in scenario 2: (a) position RMSE; (b) Euclidean distance of normalized Fourier descriptors; (c) class probability

JTC: joint tracking and classification; RHM: random hypersurface model; RMSE: root mean square error. References to color refer to the online version of this figure

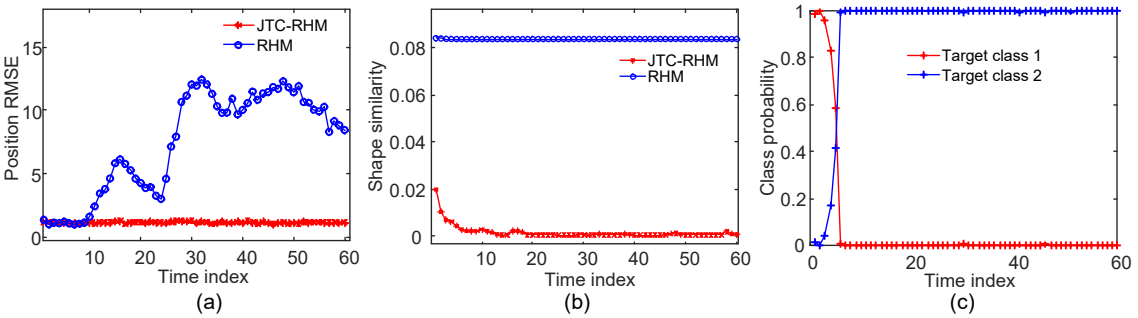

Fig. 9 Simulation results for target B in scenario 2: (a) position RMSE; (b) Euclidean distance of normalized Fourier descriptors; (c) class probability

JTC: joint tracking and classification; RHM: random hypersurface model; RMSE: root mean square error. References to color refer to the online version of this figure

proposed JTC-RHM method can significantly improve the results of extent state estimation, especially when the target is maneuvering (Figs. 8b and 9b). The proposed JTC-RHM method can achieve an accurate target classification (Figs. 8c and 9c).

5.2.3 Results for simulation scenario 3

In scenario 3, the total time duration is k=70. The target appears at time k=5, and disappears at time k=65. Figs. 10 and 11 provide simulation results for a single run in the presence of clutter and detection uncertainty. Results show that the proposed JTC-RHM-Ber filter can obtain good performance in target detection and target state estimation (both in the kinematic state and extent state).

Simulation results for the proposed JTC-RHM-Ber filter over 100 Monte Carlo trials are shown in Figs. 12 and 13. Since the proposed JTC-RHM-Ber filter can detect whether a target is present in the scene or not, we provide the estimated target cardinality and use the optimal subpattern assignment (OSPA) distance (Schuhmacher et al., 2008) to evaluate the performance of kinematic state estimation. If the target appears in the scene, we still use the Euclidean distance of the normalized Fourier descriptors to evaluate the extent state estimation, as the proposed JTC-RHM method does.

As demonstrated by Figs. 12 and 13, the proposed JTC-RHM-Ber filter can estimate the target number, target state, and target class correctly.

5.3 Algorithm complexity analysis

We analyze only the time complexity of the proposed JTC-RHM method in this study as an example. We compare the proposed JTC-RHM method with the RHM method. The time complexity analysis of the proposed JTC-RHM-Ber filter can be done in a similar way.

The complexity of the PHD filter is $\mathcal{O}(mn)$ (Mahler, 2014), where m and n are the numbers of measurements and targets, respectively. We adopt the expression of time complexity similar to that of the

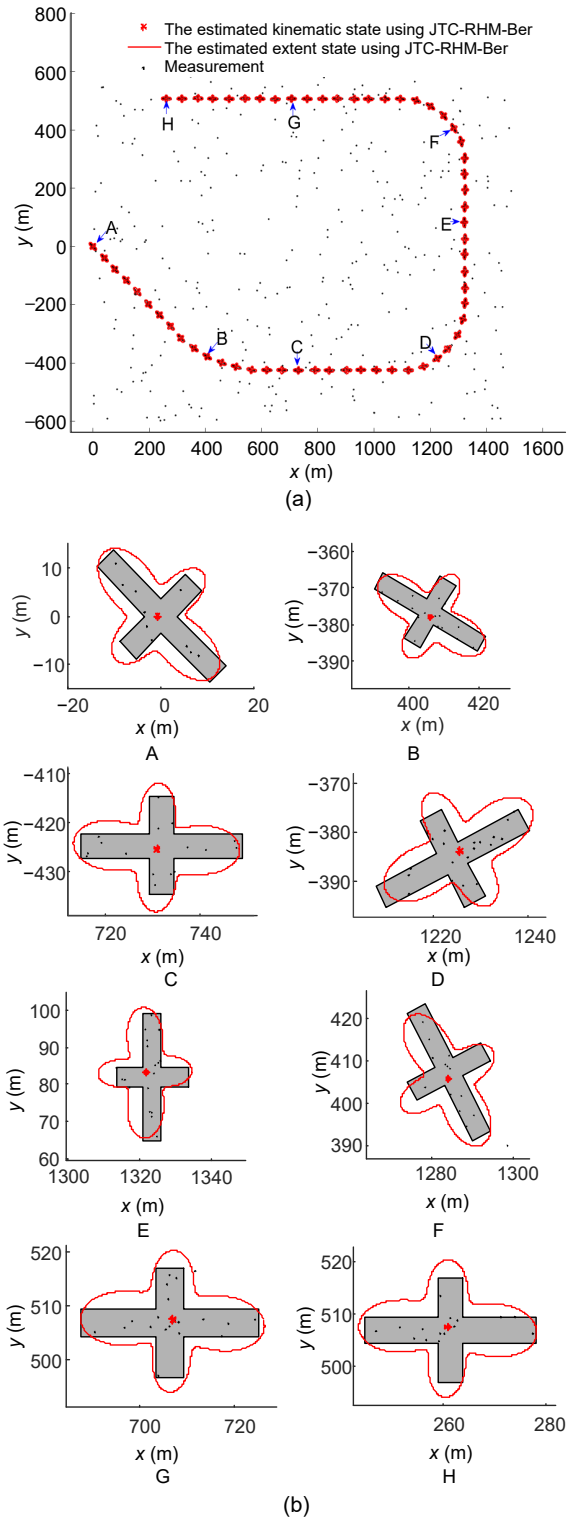

Fig. 10 Simulation results of the JTC-RHM-Ber filter for target A in scenario 3: (a) total result for a single run; (b) partially enlarged subfigures of (a)

JTC: joint tracking and classification; RHM: random hypersurface model; Ber: Bernoulli. References to color refer to the online version of this figure

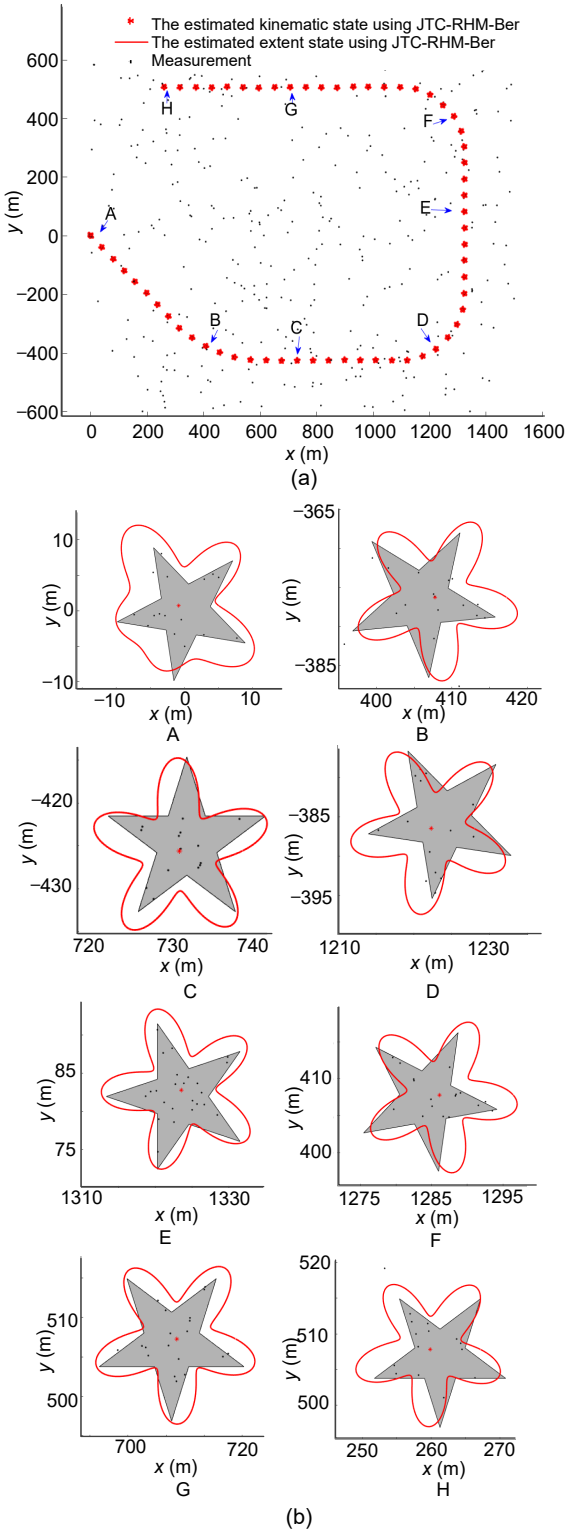

Fig. 11 Simulation results of the JTC-RHM-Ber filter for target B in scenario 3: (a) total result for a single run; (b) partially enlarged subfigures of (a)

JTC: joint tracking and classification; RHM: random hypersurface model; Ber: Bernoulli. References to color refer to the online version of this figure

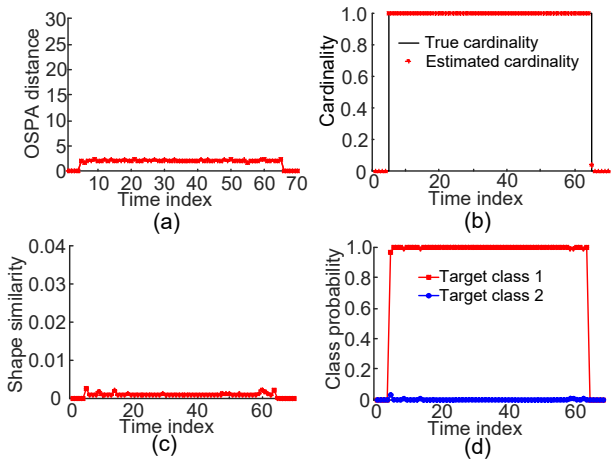

Fig. 12 Simulation results of the JTC-RHM-Ber filter for target A in scenario 3 over 100 Monte Carlo trials: (a) OSPA distance; (b) cardinality estimation; (c) shape similarity; (d) class probability

JTC: joint tracking and classification; RHM: random hypersurface model; Ber: Bernoulli; OSPA: optimal subpattern assignment. References to color refer to the online version of this figure

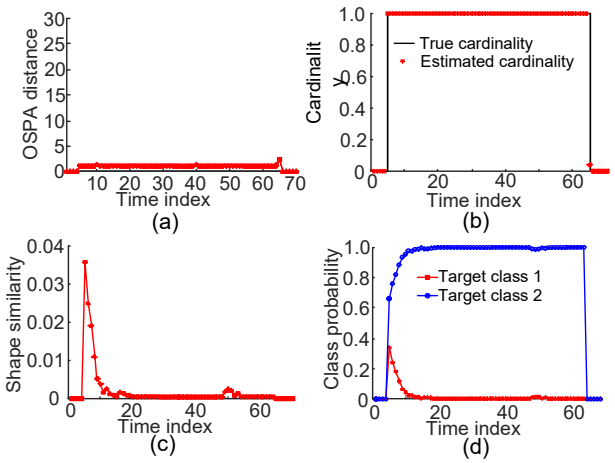

Fig. 13 Simulation results of the JTC-RHM-Ber filter for target B in scenario 3 over 100 Monte Carlo trials: (a) OSPA distance; (b) cardinality estimation; (c) shape similarity; (d) class probability

JTC: joint tracking and classification; RHM: random hypersurface model; Ber: Bernoulli; OSPA: optimal subpattern assignment. References to color refer to the online version of this figure

PHD filter. For the RHM method, all received measurements are processed sequentially. The number of received measurements at time k is $n_{k,z}$, and thus the algorithm complexity of the RHM method is $\mathcal{O}_{\mathrm{RHM}}(n_{k,z})$. The number of prior class-dependent measurements for target class c is $n^{c,\mathrm{p}}$. The proposed JTC-RHM method also uses n_c class-dependent RHM methods, with a corresponding algorithm complexity of $\mathcal{O}_{\mathrm{JTC-RHM}}\left(n_c n_{k,z}+\sum_{c=1}^{n_c} n^{c,\mathrm{p}}\right)$.

Simulations in this study are conducted using the MATLAB R2018a platform, and all experiments are performed on a computer with 3.60 GHz Intel® Core™ i7-7700 8 core CPU and 16 GB of RAM. With measurement rate λ_c=15, we conduct 100 Monte Carlo trials. The running time of the proposed JTC-RHM method and RHM method for target class A is 2.79 s and 0.74 s, respectively. For target class B, these values are 2.82 s and 0.75 s, respectively.

In this study, the number of target class is n_c=2, and the numbers of prior class-dependent measurements for the two class targets are $n^{1,\mathrm{p}}$=24 and $n^{2,\mathrm{p}}$=20, respectively. Fifteen measurements are processed sequentially for the RHM method, whereas 2×15+24+20=74 measurements need to be processed sequentially for the proposed JTC-RHM method. Theoretically, the running time ratio of the two methods is 74/15≈5. In the simulations, the actual time-consuming ratio is 2.79/0.74≈3.8 or 2.82/0.75≈3.8. The most computationally expensive operation of UKF is the unscented transformation. We model the target state using two vectors herein, and the unscented transformation is used only for the extent state. Thereby, the actual ratio of running time is less than that of the theoretical one. This further indicates the advantage of our state modeling method.

6 Conclusions

Due to the inherent relationship between target state estimation and class decision, target tracking and target classification should be processed simultaneously; i.e., joint tracking and classification needs to be performed. As the resolution of the sensor increases, a target may occupy multiple measurement cells and can be classified using the spatial extent state as feature information. When the target has a complex shape, it is not appropriate to model the target shape with a simple geometry (such as an elliptical shape). We used the star-convex RHM to model the extent state, and the class-dependent feature measurements were used as the prior information. Due to the high dimensionality of the state vector and the severe nonlinearity of the observation model, it is difficult to

obtain an analytical solution in class probability calculation. Therefore, the target state was modeled by two vectors, and the modified Euclidean distance metric of the normalized Fourier descriptors was used to calculate the class probability. Simulation results showed that the proposed JTC-RHM method outperforms the JTC-RMM method when targets have similar sizes. Compared with the conventional star-convex RHM based ETT method, the proposed JTC-RHM method can improve the estimation in both kinematic and extent states, whereas the JTC-RMM method can improve only the extent state estimation.

To solve the problem of detection uncertainty and clutter, we integrated the JTC-RHM method into the Bernoulli filter framework to obtain the JTC-RHM-Ber filter. Experimental results showed that the proposed JTC-RHM-Ber filter can not only accurately estimate the target number and target state, but also classify the target correctly.

Contributors

Liping WANG designed the research and drafted the manucript. Ronghui ZHAN helped organize the manuscript. Yuan HUANG, Jun ZHANG, and Zhaowen ZHUANG revised and finalized the paper.

Compliance with ethics guidelines

Liping WANG, Ronghui ZHAN, Yuan HUANG, Jun ZHANG, and Zhaowen ZHUANG declare that they have no conflict of interest.

References

Angelova D, Mihaylova L, 2006. Joint target tracking and classification with particle filtering and mixture Kalman filtering using kinematic radar information. *Dig Signal Process*, 16(2):180-204. https://doi.org/10.1016/j.dsp.2005.04.007

Angelova D, Mihaylova L, Petrov N, et al., 2013. A convolution particle filtering approach for tracking elliptical extended objects. Proc 16th Int Conf on Information Fusion, p.1542-1549.

Baum M, Hanebeck UD, 2014. Extended object tracking with random hypersurface models. *IEEE Trans Aerosp Electron Syst*, 50(1):149-159. https://doi.org/10.1109/taes.2013.120107

Baum M, Klumpp V, Hanebeck UD, 2010. A novel Bayesian method for fitting a circle to noisy points. Proc 13th Int Conf on Information Fusion, p.1-6. https://doi.org/10.1109/ICIF.2010.5711884

Baum M, Faion F, Hanebeck UD, 2012. Modeling the target extent with multiplicative noise. Proc 15th Int Conf on Information Fusion, p.2406-2412.

Beard M, Reuter S, Granström K, et al., 2016. Multiple extended target tracking with labeled random finite sets. *IEEE Trans Signal Process*, 64(7):1638-1653. https://doi.org/10.1109/tsp.2015.2505683

Cao W, Lan J, Li XR, 2016. Conditional joint decision and estimation with application to joint tracking and classification. *IEEE Trans Syst Man Cybern Syst*, 46(4):459-471. https://doi.org/10.1109/tsmc.2015.2442219

Cao W, Lan J, Li XR, 2018. Extended object tracking and classification using radar and ESM sensor data. *IEEE Signal Process Lett*, 25(1):90-94. https://doi.org/10.1109/lsp.2017.2757920

de Freitas A, Mihaylova L, Gning A, et al., 2019. A box particle filter method for tracking multiple extended objects. *IEEE Trans Aerosp Electron Syst*, 55(4):1640-1655. https://doi.org/10.1109/taes.2018.2874147

Eryildirim A, Guldogan MB, 2016. A Bernoulli filter for extended target tracking using random matrices in a UWB sensor network. *IEEE Sens J*, 16(11):4362-4373. https://doi.org/10.1109/jsen.2016.2544807

Feldmann M, Fränken D, Koch W, 2011. Tracking of extended objects and group targets using random matrices. *IEEE Trans Signal Process*, 59(4):1409-1420. https://doi.org/10.1109/tsp.2010.2101064

Gilholm K, Salmond D, 2005. Spatial distribution model for tracking extended objects. *IEE Proc Radar Sonar Navig*, 152(5):364-371. https://doi.org/10.1049/ip-rsn:20045114

Granström K, Lundquist C, Orguner O, 2012. Extended target tracking using a Gaussian-mixture PHD filter. *IEEE Trans Aerosp Electron Syst*, 48(4):3268-3286. https://doi.org/10.1109/taes.2012.6324703

Granström K, Reuter S, Meissner D, et al., 2014. A multiple model PHD approach to tracking of cars under an assumed rectangular shape. Proc 17th Int Conf on Information Fusion, p.1-8.

Granström K, Willett P, Bar-Shalom Y, 2015. An extended target tracking model with multiple random matrices and unified kinematics. Proc 18th Int Conf on Information Fusion, p.1007-1014.

Granström K, Baum M, Reuter S, 2017. Extended object tracking: introduction, overview, and applications. *J Adv Inform Fus*, 12(2):139-174.

Hirscher T, Scheel A, Reuter S, et al., 2016. Multiple extended object tracking using Gaussian processes. Proc 19th Int Conf on Information Fusion, p.1-8.

Hu Q, Ji HB, Zhang YQ, 2018. A standard PHD filter for joint tracking and classification of maneuvering extended targets using random matrix. *Signal Process*, 144:352-363. https://doi.org/10.1016/j.sigpro.2017.10.026

Jiang H, Zhan K, Xu L, 2015. Joint tracking and classification with constraints and reassignment by radar and ESM. *Dig Signal Process*, 40:213-223. https://doi.org/10.1016/j.dsp.2015.01.004

Knill C, Scheel A, Dietmayer K, 2016. A direct scattering model for tracking vehicles with high-resolution radars.

Proc IEEE Intelligent Vehicles Symp, p.298-303. https://doi.org/10.1109/IVS.2016.7535401

Koch JW, 2008. Bayesian approach to extended object and cluster tracking using random matrices. *IEEE Trans Aerosp Electron Syst*, 44(3):1042-1059. https://doi.org/10.1109/taes.2008.4655362

Lan J, Li XR, 2013. Joint tracking and classification of extended object using random matrix. Proc 16th Int Conf on Information Fusion, p.1550-1557.

Lan J, Li XR, 2014. Tracking of maneuvering non-ellipsoidal extended object or target group using random matrix. *IEEE Trans Signal Process*, 62(9):2450-2463. https://doi.org/10.1109/tsp.2014.2309561

Lan J, Li XR, 2016. Tracking of extended object or target group using random matrix: new model and approach. *IEEE Trans Aerosp Electron Syst*, 52(6):2973-2989. https://doi.org/10.1109/taes.2016.130346

Magnant C, Kemkemian S, Zimmer L, 2018. Joint tracking fand classification for extended targets in maritime surveillance. Proc IEEE Radar Conf, p.1117-1122. https://doi.org/10.1109/RADAR.2018.8378718

Mahler RPS, 2007. Statistical Multisource-Multitarget Information Fusion. Artech House, Boston, USA.

Mahler RPS, 2014. Advances in Statistical Multisource-Multitarget Information Fusion. Artech House, Boston, USA.

Mihaylova L, Carmi AY, Septier F, et al., 2014. Overview of Bayesian sequential Monte Carlo methods for group and extended object tracking. *Dig Signal Process*, 25:1-16. https://doi.org/10.1016/j.dsp.2013.11.006

Ristic B, Gordon N, Bessell A, 2004. On target classification using kinematic data. *Inform Fus*, 5(1):15-21. https://doi.org/10.1016/j.inffus.2003.08.002

Ristic B, Vo BT, Vo BN, et al., 2013. A tutorial on Bernoulli filters: theory, implementation and applications. *IEEE Trans Signal Process*, 61(13):3406-3430. https://doi.org/10.1109/tsp.2013.2257765

Schuhmacher D, Vo BT, Vo BN, 2008. A consistent metric for performance evaluation of multi-object filters. *IEEE Trans Signal Process*, 56(8):3447-3457. https://doi.org/10.1109/tsp.2008.920469

Sun LF, Lan J, Li XR, 2018. Joint tracking and classification of extended object based on support functions. *IET Radar Sonar Navig*, 12(7):685-693. https://doi.org/10.1049/iet-rsn.2017.0499

Wahlström N, Özkan E, 2015. Extended target tracking using Gaussian processes. *IEEE Trans Signal Process*, 63(16): 4165-4178. https://doi.org/10.1109/tsp.2015.2424194

Yang SS, Baum M, 2016. Second-order extended Kalman filter for extended object and group tracking. Proc 19th Int Conf on Information Fusion, p.1-7.

Yang SS, Baum M, 2017. Extended Kalman filter for extended object tracking. Proc IEEE Int Conf on Acoustics, Speech and Signal Processing, p.4386-4390. https://doi.org/10.1109/ICASSP.2017.7952985

Zhao YJ, Belkasim S, 2012. Multiresolution Fourier descriptors for multiresolution shape analysis. *IEEE Signal Process Lett*, 19(10):692-695. https://doi.org/10.1109/lsp.2012.2210040

Appendix: Main steps of the proposed JTC-RHM-Ber filter

Algorithm A1 Main steps of the proposed JTC-RHM-Ber filter

Input: $\left\{w_{k-1}^{(j)}, \boldsymbol{m}_{k-1}^{\mathrm{k},(j)}, \tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k},(j)}, v_{k-1}^{(j)}, \boldsymbol{V}_{k-1}^{(j)}, \left\{\boldsymbol{m}_{k-1}^{\mathrm{e},c,(j)}, \boldsymbol{P}_{k-1}^{\mathrm{e},c,(j)}, \tilde{\mu}_{k-1}^{c,(j)}\right\}_{c=1}^{n_\mathrm{c}}\right\}_{j=1}^{J_{k-1}}$, target existence probability q_{k-1}, received measurements $\mathcal{Z}_k$, and prior class-dependent measurements $\mathcal{Z}^{\mathrm{c}}$

Step 1: birth target prediction

$q_{k|k-1}=p_\mathrm{B}(1-q_{k-1})+p_\mathrm{S}q_{k-1}$;
$l=0$;
for $j=1, 2, \ldots, J_{k,\mathrm{B}}$
  $l=l+1$;
  $w_{k|k-1}^{(l)}=\dfrac{p_\mathrm{B}(1-q_{k-1})}{q_{k|k-1}}w_{k-1,\mathrm{B}}^{(j)}$;
  $\boldsymbol{m}_{k|k-1}^{\mathrm{k},(l)}=\boldsymbol{m}_{k,\mathrm{B}}^{\mathrm{k},(j)}$, $\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(l)}=\tilde{\boldsymbol{P}}_{k,\mathrm{B}}^{\mathrm{k},(j)}$;
  $v_{k|k-1}^{(l)}=v_{k,\mathrm{B}}^{(j)}$, $\boldsymbol{V}_{k|k-1}^{\mathrm{k},(l)}=\boldsymbol{V}_{k,\mathrm{B}}^{\mathrm{k},(j)}$;
  for $c=1, 2, \ldots, n_c$
    $\boldsymbol{m}_{k|k-1}^{\mathrm{e},c,(l)}=\boldsymbol{m}_{k,\mathrm{B}}^{\mathrm{e},c,(j)}$, $\boldsymbol{P}_{k|k-1}^{\mathrm{e},c,(l)}=\boldsymbol{P}_{k,\mathrm{B}}^{\mathrm{e},c,(j)}$, $\tilde{\mu}_{k|k-1}^{c,(l)}=\tilde{\mu}_{k,\mathrm{B}}^{c,(j)}$;
  end
end

To be continued

Algorithm A1

Step 2: surviving target prediction

for j=1, 2, …, J_{k-1}

l=l+1;

$\boldsymbol{m}_{k|k-1}^{\mathrm{k},(l)}=\boldsymbol{F}_k^{\mathrm{k}}\boldsymbol{m}_{k-1}^{\mathrm{k},(j)}$, $\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(l)}=\tilde{\boldsymbol{F}}_k^{\mathrm{k}}\tilde{\boldsymbol{P}}_{k-1}^{\mathrm{k},(j)}(\tilde{\boldsymbol{F}}_k^{\mathrm{k}})^{\mathrm{T}}+\tilde{\boldsymbol{Q}}_k^{\mathrm{k}}$;

$\lambda_{k-1}^{(j)}=v_{k-1}^{(j)}-2d-2$;

$$v_{k|k-1}^{(l)}=\frac{2\delta_k\left(\lambda_{k-1}^{(j)}+1\right)\left(\lambda_{k-1}^{(j)}-1\right)\left(\lambda_{k-1}^{(j)}-2\right)}{\left(\lambda_{k-1}^{(j)}\right)^2\left(\lambda_{k-1}^{(j)}+\delta_k\right)}+2d+4;$$

$$V_{k|k-1}^{(l)}=\frac{\delta_k\left(v_{k|k-1,\mathrm{S}}^{(j)}-2d-2\right)}{\lambda_{k-1}^{(j)}}\boldsymbol{A}_k\boldsymbol{V}_{k-1}^{(j)}\boldsymbol{A}_k^{\mathrm{T}};$$

for c=1, 2, …, n_c

$\boldsymbol{m}_{k|k-1}^{\mathrm{e},c,(l)}=\boldsymbol{m}_0^{\mathrm{e},c}$, $\boldsymbol{P}_{k|k-1}^{\mathrm{e},c,(l)}=\boldsymbol{P}_0^{\mathrm{e},i}$, $\tilde{\mu}_{k|k-1}^{c,(l)}=\tilde{\mu}_{k-1}^{c,(j)}$;

end

end

$J_{k|k-1}$=l.

Step 3: non-detection part update

for j=1, 2, …, $J_{k|k-1}$

$w_k^{(j)}=(1-p_{\mathrm{D}})^{l_k}w_{k|k-1}^{(j)}$;

$\boldsymbol{m}_k^{\mathrm{k},(j)}=\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}$, $\tilde{\boldsymbol{P}}_k^{\mathrm{k},(j)}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(j)}$, $v_k^{(j)}=v_{k|k-1}^{(j)}$, $\boldsymbol{V}_k^{(j)}=\boldsymbol{V}_{k|k-1}^{(j)}$;

for c=1, 2, …, n_c

$\boldsymbol{m}_k^{\mathrm{e},c,(j)}=\boldsymbol{m}_{k|k-1}^{\mathrm{e},c,(j)}$, $\boldsymbol{P}_k^{\mathrm{e},c,(j)}=\boldsymbol{P}_{k|k-1}^{\mathrm{e},c,(j)}$, $\tilde{\mu}_k^{c,(l)}=\tilde{\mu}_{k|k-1}^{c,(j)}$;

end

end

Step 4: detection part update

Compute measurement set $\{\mathcal{P}_p\}_{p=1}^{n_{\mathrm{p}}}$ and l_k, where $\mathcal{P}_p\triangleq\{\mathcal{W}_\omega^p\}_{\omega=1}^{n_\omega}$ (Granström et al., 2012; Eryildirim and Guldogan, 2016);

l=1, Δ_k=0, $\Delta_{k,\mathrm{temp}}$=0;

for p=1, 2, …, n_{p}

for ω=1, 2, …, n_ω

$$\psi_k=\frac{l_k!}{(l_k-|\mathcal{W}_\omega^p|)!}\frac{p_{\mathrm{D}}^{|\mathcal{W}_\omega^p|}}{(1-p_{\mathrm{D}})^{|\mathcal{W}_\omega^p|-l_k}};$$

l=l+1;

for j=1, 2, …, $J_{k|k-1}$

$$\overline{\boldsymbol{z}}_k=\frac{1}{|\mathcal{W}_\omega^p|}\sum\nolimits_{\boldsymbol{z}_k\in\mathcal{W}_\omega^p}\boldsymbol{z}_k,\ \overline{\boldsymbol{Z}}_k=\sum\nolimits_{\boldsymbol{z}_k\in\boldsymbol{W}_\omega^p}(\boldsymbol{z}_k-\overline{\boldsymbol{z}}_k)(\boldsymbol{z}_k-\overline{\boldsymbol{z}}_k)^{\mathrm{T}};$$

$$\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}}=\tilde{\boldsymbol{H}}\tilde{\boldsymbol{P}}_{k|k-1}^{(j)}\tilde{\boldsymbol{H}}^{\mathrm{T}}+\frac{[\det(\boldsymbol{B}_k)]^{2/d}}{|\mathcal{W}_\omega^p|},\ \boldsymbol{N}_k=\left(\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}}\right)^{-1}\left(\overline{\boldsymbol{z}}_k-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}\right)\left(\overline{\boldsymbol{z}}_k-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}\right)^{\mathrm{T}};$$

$$\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}=\tilde{\boldsymbol{P}}_{k|k-1}^{\mathrm{k},(j)}\tilde{\boldsymbol{H}}^{\mathrm{T}}\left(\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}}\right)^{-1},\ \boldsymbol{K}_{k|k-1}^{\mathrm{k}}=\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}\otimes\boldsymbol{I}_2;$$

$$\boldsymbol{m}_k^{\mathrm{k},(lJ_{k|k-1}+j)}=\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}+\boldsymbol{K}_{k|k-1}^{\mathrm{k}}\left(\overline{\boldsymbol{z}}_k-\boldsymbol{H}\boldsymbol{m}_{k|k-1}^{\mathrm{k},(j)}\right);$$

$$\tilde{\boldsymbol{P}}_k^{\mathrm{k},(lJ_{k|k-1}+j)}=\boldsymbol{P}_{k|k-1}^{\mathrm{k},(j)}-\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}}\left(\tilde{\boldsymbol{K}}_{k|k-1}^{\mathrm{k}}\right)^{\mathrm{T}};$$

$$v_k^{(lJ_{k|k-1}+j)}=|\mathcal{W}_\omega^p|+v_{k|k-1}^{(j)},\ \boldsymbol{V}_k^{(lJ_{k|k-1}+j)}=\boldsymbol{V}_{k|k-1}^{(j)}+\boldsymbol{B}_k^{-1}\overline{\boldsymbol{Z}}_k\boldsymbol{B}_k^{-\mathrm{T}}+\boldsymbol{N}_k;$$

$$\mathcal{L}_k^{(j,\mathcal{W})}=\left[|\boldsymbol{B}_k|^{\frac{-(|\mathcal{W}_\omega^p|-1)}{2}}[\det(\boldsymbol{V}_{k|k-1}^{(j)})]^{\frac{v_{k|k-1}^{(j)}-d-1}{2}}\Gamma_d\left(\frac{v_k^{(lJ_{k|k-1}+j)}-d-1}{2}\right)\right]$$

To be continued

Algorithm A1

$$\cdot\left[\left((\pi)^{-|\mathcal{W}_\omega^p|}\left|\mathcal{W}_\omega^p\right|\det(\tilde{\boldsymbol{S}}_{k|k-1}^{\mathrm{k}})\right)^{\frac{d}{2}}\left[\det\left(\boldsymbol{V}_k^{(l\mathcal{J}_{k|k-1}+j)}\right)\right]^{\frac{v_k^{(j,\mathcal{W})}-d-1}{2}}\Gamma_d\left(\frac{v_{k|k-1}^{(j)}-d-1}{2}\right)\right]^{-1};$$

$$w_k^{(lJ_{k|k-1}+j)}=\psi_k\left(\lambda_{\mathrm{c}}\mathcal{C}(\boldsymbol{z})\right)^{-|\mathcal{W}|}\mathcal{L}_k^{(j,\mathcal{W})}w_{k|k-1}^{(j)},\ \Delta_{k,\mathrm{temp}}=\Delta_{k,\mathrm{temp}}+w_k^{(lJ_{k|k-1}+j)};$$

for c=1, 2, …, n_c

$$\theta_k^{\mathrm{k}}=\arctan\left(\frac{\boldsymbol{m}_k^{\mathrm{k},(lJ_{k|k-1}+j)}(4)}{\boldsymbol{m}_k^{\mathrm{k},(lJ_{k|k-1}+j)}(3)}\right),\ \mathcal{Z}_k^c=\begin{bmatrix}\cos(\theta_k^{\mathrm{k}}) & -\sin(\theta_k^{\mathrm{k}})\\ \sin(\theta_k^{\mathrm{k}}) & \cos(\theta_k^{\mathrm{k}})\end{bmatrix}\mathcal{Z}^c;$$

for $m=1,2,\cdots,|\mathcal{Z}_k^c|,\cdots,|\mathcal{Z}_k^c|+|\mathcal{W}_\omega^p|$

$$\boldsymbol{m}_k^{\mathrm{p}}=[\boldsymbol{m}_k^{\mathrm{k},(lJ_{k|k-1}+j)}(1),\ \boldsymbol{m}_k^{\mathrm{k},(lJ_{k|k-1}+j)}(2)]^{\mathrm{T}};$$

$$\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c}=[(\boldsymbol{m}_k^{\mathrm{p}})^{\mathrm{T}},(\boldsymbol{m}_{k|k-1}^{\mathrm{e},c})^{\mathrm{T}}]^{\mathrm{T}},\ \boldsymbol{P}_{k|k-1}^{\mathrm{pe},c}=\mathrm{diag}(\boldsymbol{I}_2,\boldsymbol{P}_{k|k-1}^{\mathrm{e},c});$$

$$\boldsymbol{m}_{k|k-1}^{\mathrm{aug},c}=[(\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c})^{\mathrm{T}},\mu_s,\boldsymbol{0}_{2\times1}],\ \boldsymbol{P}_{k|k-1}^{\mathrm{aug},c}=\mathrm{diag}(\boldsymbol{P}_{k|k-1}^{\mathrm{pe},c},\sigma_s^2,\boldsymbol{R}_k);$$

$$\left(\left\{\tilde{\boldsymbol{m}}_{k|k-1,i}^{\mathrm{aug},c}\right\}_{i=0}^{2n^{\mathrm{aug}}},\left\{\tilde{\omega}_{k|k-1,i}^{\mathrm{m},c}\right\}_{i=0}^{2n^{\mathrm{aug}}},\left\{\tilde{\omega}_{k|k-1,v}^{\mathrm{P},c}\right\}_{i=0}^{2n^{\mathrm{aug}}}\right)\leftarrow\mathrm{UT}(\boldsymbol{m}_{k|k-1}^{\mathrm{aug},c},\boldsymbol{P}_{k|k-1}^{\mathrm{aug},c});$$

$$\tilde{\boldsymbol{y}}_{k|k-1,i}=h^*(\tilde{\boldsymbol{m}}_{k|k-1,i}^{c,\mathrm{aug}}(1:n^{\mathrm{aug}}-3),\boldsymbol{v}_k,s_k,\boldsymbol{z}_m);$$

$$\hat{\boldsymbol{z}}_{k|k-1}=\sum_{i=1}^{2n^{\mathrm{aug}}}\tilde{\omega}_{k|k-1,v}^{\mathrm{m},c}\tilde{\boldsymbol{y}}_{k|k-1,i},\ \boldsymbol{P}_{\boldsymbol{y}_k\boldsymbol{y}_k}=\sum_{i=0}^{2n^{\mathrm{aug}}}\tilde{\omega}_{k|k-1,i}^{\mathrm{P},c}[\tilde{\boldsymbol{y}}_{k|k-1,i}-\hat{\boldsymbol{z}}_{k|k-1}][\tilde{\boldsymbol{y}}_{k|k-1,i}-\hat{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}};$$

$$\boldsymbol{P}_{\boldsymbol{x}_k\boldsymbol{y}_k}=\sum_{i=0}^{2n^{\mathrm{aug}}}\tilde{\omega}_{k|k-1,i}^{\mathrm{P},c}[\tilde{\boldsymbol{m}}_{k|k-1,i}^{c,\mathrm{aug}}-\boldsymbol{m}_{k|k-1}^{\mathrm{aug,c}}][\tilde{\boldsymbol{y}}_{k|k-1,i}-\hat{\boldsymbol{z}}_{k|k-1}]^{\mathrm{T}},\ \boldsymbol{K}_k^{\mathrm{pe}}=\boldsymbol{P}_{\boldsymbol{x}_k\boldsymbol{y}_k}\boldsymbol{P}_{\boldsymbol{y}_k\boldsymbol{y}_k}^{-1};$$

$$\boldsymbol{m}_k^{\mathrm{pe},c}=\boldsymbol{m}_{k|k-1}^{\mathrm{pe},c}+\boldsymbol{K}_k^{\mathrm{pe}}(0-\hat{\boldsymbol{z}}_{k|k-1}),\ \boldsymbol{P}_k^{\mathrm{pe},c}=\boldsymbol{P}_{k|k-1}^{\mathrm{pe},c}-\boldsymbol{K}_k^{\mathrm{pe}}\boldsymbol{P}_{\boldsymbol{y}_k\boldsymbol{y}_k}(\boldsymbol{K}_k^{\mathrm{pe}})^{\mathrm{T}};$$

end

$$\boldsymbol{m}_k^{\mathrm{e},c,(lJ_{k|k-1}+j)}=\boldsymbol{m}_k^{\mathrm{pe,c}}(3:\mathrm{end}),\ \boldsymbol{P}_k^{\mathrm{e},c,(J_{k|k-1}+j)}=\boldsymbol{P}_k^{\mathrm{pe},c}(3:\mathrm{end},3:\mathrm{end});$$

$$\tilde{\mu}_k^{c,(l\mathcal{J}_{k|k-1}+j)}=\tilde{\eta}_k^c\tilde{\mu}_{k|k-1}^{c,(j)},\ \tilde{\eta}_k^c\ \text{is obtained by Eq. (16)};$$

end

$$\tilde{\mu}_k^{c,(lJ_{k|k-1}+j)}=\frac{\tilde{\mu}_k^{c,(lJ_{k|k-1}+j)}}{\sum_{c=1}^{n_c}\tilde{\mu}_k^{c,(lJ_{k|k-1}+j)}};$$

end

end

end

$J_k=lJ_{k|k-1}+J_{k|k-1}$, $\Delta_k=1-(1-p_{\mathrm{D}})^{l_k}-\Delta_{k,\mathrm{temp}}$;

for j=1, 2, …, J_k

$$w_k^{(j)}=\frac{w_k^{(j)}}{1-\Delta_k};$$

end

$$q_k=\frac{1-\Delta_k}{1-q_{k|k-1}\Delta_k}q_{k|k-1}.$$

Step 5: merging and pruning (Granström et al., 2012)

Step 6: state extraction

if $q_k\geq0.5$

$$j^*=\arg\max_j w_k^{(j)};$$

$$\hat{\underline{\boldsymbol{m}}}_k=\left(\boldsymbol{m}_k^{\mathrm{k},(j^*)},\ \sum_{c=1}^{n_c}\boldsymbol{m}_k^{\mathrm{e},(j^*)}\tilde{\mu}_k^{c,(j^*)}\right);$$

end

Output: $\hat{\underline{\boldsymbol{m}}}_k$ and $\tilde{\mu}_k^{c,(j^*)}$