

Article

Variational Bayesian Based Adaptive Shifted Rayleigh Filter for Bearings-Only Tracking in Clutters

Jing Hou , Yan Yang * and Tian Gao

School of Electronic and Information, Northwestern Polytechnical University, Xi'an 710072, China; jhou0825@nwpu.edu.cn (J.H.); tiangao@nwpu.edu.cn (T.G.)

* Correspondence: yangyan7003@nwpu.edu.cn

Received: 11 February 2019; Accepted: 26 March 2019; Published: 28 March 2019



Abstract: This paper considers bearings-only target tracking in clutters with uncertain clutter probability. The traditional shifted Rayleigh filter (SRF), which assumes known clutter probability, may have degraded performance in challenging scenarios. To improve the tracking performance, a variational Bayesian-based adaptive shifted Rayleigh filter (VB-SRF) is proposed in this paper. The target state and the clutter probability are jointly estimated to account for the uncertainty in clutter probability. Performance of the proposed filter is evaluated by comparing with SRF and the probability data association (PDA)-based filters in two scenarios. Simulation results show that the proposed VB-SRF algorithm outperforms the traditional SRF and PDA-based filters especially in complex adverse scenarios in terms of track continuity, track accuracy and robustness with a little higher computation complexity.

Keywords: bearings-only tracking; clutter; variational Bayesian; Shifted Rayleigh Filter

1. Introduction

Bearings-only target tracking is to estimate the current position and velocity of a target using only the noise-corrupted bearing measurements from one or multiple observer platforms. It is an important tracking problem that arises in both military and civilian applications, such as underwater sonar tracking, bistatic radar, air traffic control and computer vision [1].

Because of the intrinsic nonlinearities in the measurement models, it is difficult to acquire an optimal solution of this problem. Several suboptimal algorithms have been developed for bearings-only tracking in the literature. The extended Kalman filter (EKF) [2] in the Cartesian coordinate system is an early attempt. However, it is easy to diverge. To improve the stability of the EKF, the bearings-only tracking problem was formulated in modified polar coordinates, resulting in the modified polar coordinate EKF (MPEKF) [3]. However, it requires good initialization to guarantee convergence. The well-known pseudo-linear estimator (PLE) [4] was also developed to solve the bearings-only tracking problem. However, it gives a biased estimate at long ranges. In recent years, some bias compensation techniques were developed to improve the performance of PLE [5–7]. In addition, more sophisticated nonlinear Kalman filtering algorithms, such as unscented Kalman filter (UKF) [8,9], cubature Kalman filter (CKF) [1] and particle filter (PF) [10,11] were applied for bearings-only target tracking. PF can provide good performance but at the price of heavy computation load. Noteworthily, Clark et al. [12,13] proposed a novel shifted Rayleigh filter (SRF) for bearings-only tracking, which is still based on the approximation of conditional expectations but with novel feature which is performing a calculation to exploit the essential structure of the nonlinearities in a new way. It is shown to exhibit similar performance to the PF in certain challenging scenarios with much lower computational complexity [14].

However, these algorithms do not consider the impact of clutter which makes the bearings-only tracking problem more intractable. Apparently, the classical treatment of clutter in target tracking

problem can be extended to the bearings-only tracking problem. For example, Reference [15] integrated the maximum entropy fuzzy probabilistic data association (MEFPDA) with the square-root cubature Kalman filter (SCKF) to deal with the clutter in bearings-only tracking. Clark et al. also included the effect of clutter measurements into the SRF algorithm in [12]. However, in the classical algorithms, the clutter probability is usually assumed known and constant, which maybe time-varying or hard to determine in advance especially in adverse scenarios. Use of incorrect value of the clutter probability may lead to track accuracy degradation even track loss. A straight-forward idea is to account for the unknown clutter probability in the process of estimation of the state. That is, we need to solve the problem of bearings-only target tracking with uncertain parameter of clutter probability.

As we know, the Bayesian approach is the most general approach of solving the problem with uncertain parameters. However, it is not trivial to get the analytical solution for most Bayesian approaches due to complex nonlinear probability density function or high dimension of integration. Recently, the variational Bayesian (VB)-based adaptive filters [16–18] have drawn extensive attention, which utilize a new simpler, analytically tractable distribution to approximate the true posterior distribution so that avoiding direct calculations of complex integrals. Its adaptive strategy has a strong ability of tracking time-varying parameters. Therefore, we adopt the VB method to jointly estimate the target state and the clutter probability within the framework of SRF in this paper for bearings-only tracking in clutters. By establishing a conjugate exponential model for clutter probability and data association indicator, the proposed filter is derived using the iterative filtering framework. The tracking performance of the proposed VB-SRF is evaluated by comparing with SRF, PDA-SCKF [15] and MEFPDA-SCKF [15] via two simulation examples. It shows that the proposed filter outperforms the traditional SRF and two PDA-SCKF-based filters in complex mismatched scenarios in terms of track continuity and track accuracy but at the cost of higher computation complexity.

The remainder of the paper is organized as follows. Section 2 gives the problem formulation. In Section 3, the variational Bayesian filtering is described. Section 4 derives the VB-based adaptive SRF. Section 5 provides simulation results and performance evaluation of the proposed approach, followed by the conclusions in Section 6.

2. The Shifted Rayleigh Filter Algorithm

2.1. The Bearing Model

Considering the shifted Rayleigh filter (SRF) for bearings-only tracking in R^2 , the state equation and the measurement equation are described as [13]:

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{u}_{k-1}^s + \mathbf{v}_{k-1} \quad (1)$$

$$\begin{aligned} \mathbf{y}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{u}_k^m + \mathbf{w}_k \\ \mathbf{b}_k &= \Pi(\mathbf{y}_k) \end{aligned} \quad (2)$$

where, \mathbf{x}_k is the state vector which describes the position and velocity of the target; \mathbf{b}_k is the noisy bearing measurement. Π denotes the projection of the plane onto the unit circle. That is taking a 2-vector \mathbf{y}_k into its normalized form $\mathbf{y}_k/|\mathbf{y}_k|$. Then $\mathbf{b}_k = (\sin\theta_k, \cos\theta_k)^T$, where θ_k is the bearing of the target position relative to the sensor platform. \mathbf{F}_{k-1} and \mathbf{H}_k are the state transition matrix and measurement matrix, respectively. \mathbf{u}_{k-1}^s and \mathbf{u}_k^m are the inputs to the system to increase the versatility of the model, for example, to reflect known perturbations to the dynamics and changes in sensor location; \mathbf{v}_{k-1} is the Gaussian process noise with zero mean and covariance \mathbf{Q}_v , and \mathbf{w}_k is the Gaussian measurement noise with zero mean and covariance \mathbf{Q}_w and independent of \mathbf{v}_k .

The unusual point in the measurement model is that the noise \mathbf{w}_k is modelled as additive noise present in an “augmented” measurement, \mathbf{y}_k , of the Cartesian coordinates of target relative to the sensor platform, which is projected onto the plane to generate the actual bearing measurement \mathbf{b}_k . It is different with the traditional “angle plus white noise” model, expressed as

$$\theta_k = \arctan(d_1/d_2) + \epsilon_k \quad (3)$$

where, d_1, d_2 are the components of the displacement vector $\mathbf{d}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{u}_k^m$. ϵ_k is the sensor noise with Gaussian distribution $\mathcal{N}(0, \sigma^2)$ and independent of the displacement vector \mathbf{d}_k .

Actually, as explained in [13], the shifted Rayleigh bearing model (2) can be related with the traditional model (3) by making a variant on the shifted Rayleigh noise model

$$\begin{aligned} \mathbf{y}'_k &= \mathbf{d}_k + \|\mathbf{d}_k\| \mathbf{e}' \\ \mathbf{b}'_k &= \Pi(\mathbf{y}'_k) \end{aligned} \quad (4)$$

where, \mathbf{e}' is an $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_{2 \times 2})$ distributed random variable. The only difference with model (2) is the noise term \mathbf{w}_k used to construct the augmented measurement \mathbf{y}_k is replaced by $\|\mathbf{d}_k\| \mathbf{e}'$. $\|\mathbf{d}_k\| \mathbf{e}'$ differs from \mathbf{w}_k , but has identical first and second moments, and is uncorrelated with \mathbf{d}_k .

The angular θ'_k of the modified vector bearing \mathbf{b}'_k can be represented as

$$\theta'_k = \arctan(d_1/d_2) + \epsilon'_k \quad (5)$$

where ϵ'_k is a zero mean random variable, restricted to $[-\pi, \pi]$, independent of \mathbf{d}_k , with density $\alpha_\sigma(\cdot)$:

$$\alpha_\sigma(\theta) = \frac{e^{-1/2\sigma^2}}{2\pi} \left(1 + \sqrt{2\pi} \frac{\cos\theta}{\sigma} F_{normal} \left(\frac{\cos\theta}{\sigma} \right) e^{1/2(\cos\theta/\sigma)^2} \right) \quad (6)$$

where, $F_{normal}(\cdot)$ is the cumulative distribution function of a standard $\mathcal{N}(0, 1)$ variable.

Note that θ' given by (5), is very close to the bearing θ in the standard model (3). The only difference is that the densities of the noise terms used in their construction are $\alpha_\sigma(\theta)$ and the normal $\mathcal{N}(0, \sigma^2)$ density, respectively. Reference [13] plots the two densities for $\sigma^2 \leq 1$. It shows the two density functions are virtually indistinguishable.

Given the bearing model, the SRF is to calculate the estimates of the conditional mean and covariance of the target state \mathbf{x}_k , given measurements up to time k , $\mathbf{b}_{1:k}$. That is,

$$\hat{\mathbf{x}}_{k|k} = E[\mathbf{x}_k | \mathbf{b}_{1:k}], \quad \mathbf{P}_{k|k} = \text{cov}[\mathbf{x}_k | \mathbf{b}_{1:k}] \quad (7)$$

The formulas for the SRF algorithm can be seen in [13].

2.2. The Treatment of Clutter

Accounting for the effects of clutter on the bearing measurements, we represent a cluttered bearing measurement as

$$z_k = (1 - r_k)\theta_k + r_k U_k \quad (8)$$

where U_k is the bearing measurement of the clutter, which is assumed to be an uniform random variable on $[-\pi, \pi]$; θ_k is the bearing measurement of the actual target from the sensor, r_k is defined as data association indicator at time k .

$$r_k = \begin{cases} 1 & \text{if the measurement is from clutter} \\ 0 & \text{if the measurement is from target} \end{cases} \quad (9)$$

The prior of r_k is assumed independent of the previous data associations and can be described as

$$p(r_k) = \begin{cases} \xi & \text{if } r_k = 1 \\ 1 - \xi & \text{if } r_k = 0 \end{cases} \quad (10)$$

Then the likelihood of the measurement is

$$p(z_k|\mathbf{x}_k, r_k) = \begin{cases} 1/2\pi & \text{if } r_k = 1 \\ f(\theta_k|\mathbf{x}_k) & \text{if } r_k = 0 \end{cases} \quad (11)$$

It can be represented in two forms:

$$p(z_k|\mathbf{x}_k, r_k) = \frac{1}{2\pi}r_k + f(\theta_k|\mathbf{x}_k)(1 - r_k) \quad (12)$$

$$= \left(\frac{1}{2\pi}\right)^{r_k} f(\theta_k|\mathbf{x}_k)^{1-r_k} \quad (13)$$

Note that the two forms are used for different purpose in the later section. $f(\theta_k|\mathbf{x}_k)$ is the likelihood of the target bearing measurement, whose expression is given in Appendix A.

Here, we need to explain the reason for using this representation (8) of the cluttered measurement. The information carried in (8) is identical with the projection measurement $\mathbf{z}_k^b = [\sin z_k, \cos z_k]$. So the mean and covariance of the target state \mathbf{x}_k conditioned on cluttered measurement \mathbf{z}_k^b , which are the aims needs to be achieved in the SRF, are equivalent with these conditioned on z_k . Furthermore, this representation (8) is more convenient for calculation. Thus, z_k instead of \mathbf{z}_k^b is adopted to represent the clutter measurement.

Based on the cluttered measurement model, given clutter probability ζ , the state estimate and covariance can be calculated as follows.

Suppose the density of \mathbf{x}_{k-1} given measurements up to $k-1$, $p_{k-1|k-1}(\mathbf{x}_{k-1})$, is normal with mean $\bar{\mathbf{x}}_{k-1|k-1}$ and covariance $\bar{\mathbf{P}}_{k-1|k-1}$. The posterior density of state \mathbf{x}_k conditioned on cluttered bearing measurements $\mathbf{z}_{1:k}$ is given as

$$\begin{aligned} p(\mathbf{x}_k|\mathbf{z}_{1:k}) &= p(\mathbf{x}_k|z_k, \mathbf{x}_{k-1} \sim p_{k-1|k-1}(\mathbf{x}_{k-1})) \\ &= q_k(0)p_{k|k}(\mathbf{x}_k) + q_k(1)p_{k|k-1}(\mathbf{x}_k) \end{aligned} \quad (14)$$

where $q_k(i) = p(r_k = i|z_k, \mathbf{z}_{1:k-1})$, $p_{k|k}(\mathbf{x}_k)$ is the non-normal density of \mathbf{x}_k conditioned on $r_k = 0$ and z_k , or, equivalently, on θ_k , and $p_{k|k-1}(\mathbf{x}_k)$ is the density of \mathbf{x}_k when there is no target measurement at time k . It is normal with mean $\hat{\mathbf{x}}_{k|k-1}$ and covariance $\mathbf{P}_{k|k-1}$. Thus, the state estimate and covariance at time k can be obtained as

$$\begin{aligned} \bar{\mathbf{x}}_{k|k} &= E[\mathbf{x}_k|z_k, \mathbf{x}_{k-1} \sim p_{k-1|k-1}(\mathbf{x}_{k-1})] \\ &= q_k(0)\hat{\mathbf{x}}_{k|k} + q_k(1)\hat{\mathbf{x}}_{k|k-1} \end{aligned} \quad (15)$$

$$\begin{aligned} \bar{\mathbf{P}}_{k|k} &= cov[\mathbf{x}_k|z_k, \mathbf{x}_{k-1} \sim p_{k-1|k-1}(\mathbf{x}_{k-1})] \\ &= q_k(0)(\mathbf{P}_{k|k} + (\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}_{k|k})(\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}_{k|k})^T) \\ &\quad + q_k(1)(\mathbf{P}_{k|k-1} + (\hat{\mathbf{x}}_{k|k-1} - \bar{\mathbf{x}}_{k|k})(\hat{\mathbf{x}}_{k|k-1} - \bar{\mathbf{x}}_{k|k})^T) \end{aligned} \quad (16)$$

where $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}$ are the state estimate and its covariance based on the actual target measurement. $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ are the predicted target state estimate and covariance. All of these can be obtained using the basic formulas of SRF.

The conditional densities $q_k(0)$ and $q_k(1)$ are given by the equations

$$q_k(0) = c_k(1 - \zeta)f_k(\theta_k|\mathbf{z}_{1:k-1}) \quad (17)$$

$$q_k(1) = \frac{c_k\zeta}{2\pi} \quad (18)$$

where c_k is the normalizing constant and $f_k(\theta_k|\mathbf{z}_{1:k-1})$ is the density of the actual target bearing θ_k conditioned on measurements $\mathbf{z}_{1:k-1}$. The expression is given in Appendix B.

However, in a complex environment, the probability of clutter is time-varying or hard to determine in advance. In this case, the probability ζ is unknown, so the above formulas are not applicable. In this paper, we resort the VB method to find the joint posterior density of \mathbf{x}_k and ζ so as to account for the uncertainty in clutter probability.

3. Variational Bayesian Filtering

In this section, we first review the conjugate exponential (CE) model, and then gives the variational Bayesian solution of the CE model.

3.1. Conjugate Exponential Model

Given measurements $\mathbf{z}_{1:k-1}$, the posterior of the system state $p(\mathbf{x}_{k-1}|\mathbf{z}_{1:k-1})$ and the posterior of the parameter $p(\mathbf{r}_{k-1}|\mathbf{z}_{1:k-1})$, we assume the complete-data likelihood in the exponential family:

$$p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{r}_k, \mathbf{z}_{1:k-1}) = g(\mathbf{r}_k)f(\mathbf{x}_k, \mathbf{z}_k)e^{\phi(\mathbf{r}_k)^T u(\mathbf{x}_k, \mathbf{z}_k)} \quad (19)$$

where, $\phi(\mathbf{r}_k)$ is the vector of natural parameters \mathbf{r}_k , u and f are known functions, and g is a normalization constant:

$$g(\mathbf{r}_k)^{-1} = \int f(\mathbf{x}_k, \mathbf{z}_k)e^{\phi(\mathbf{r}_k)^T u(\mathbf{x}_k, \mathbf{z}_k)} d\mathbf{x}_k d\mathbf{z}_k$$

The parameter prior is conjugate to the complete-data likelihood:

$$p(\mathbf{r}_k|\alpha_k^-, \beta_k^-) = h(\alpha_k^-, \beta_k^-)g(\mathbf{r}_k)^{\beta_k^-} e^{\phi(\mathbf{r}_k)^T \alpha_k^-} \quad (20)$$

where α_k^- and β_k^- are hyperparameters of the prior, and h is a normalization constant. Note the prior $p(\mathbf{r}_k|\alpha_k^-, \beta_k^-)$ is said to be conjugate to the likelihood $p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{r}_k)$ if and only if the posterior

$$p(\mathbf{r}_k|\alpha_k, \beta_k) \propto p(\mathbf{r}_k|\alpha_k^-, \beta_k^-)p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{r}_k)$$

is of the same parametric form as the prior. Then we call models that satisfy Equations (19) and (20) conjugate-exponential.

3.2. VB Approximation Method

Applying Bayes' rule, we have the joint posterior of \mathbf{x}_k and r_k as

$$p(\mathbf{x}_k, \mathbf{r}_k|\mathbf{z}_{1:k}) \propto p(\mathbf{x}_k, \mathbf{z}_k|\mathbf{r}_k, \mathbf{z}_{1:k})p(\mathbf{r}_k|\alpha_k^-, \beta_k^-) \quad (21)$$

The analytic solution to (21) would be difficult to calculate. Here, we use the VB method to approximate the true posterior distribution with a product of tractable marginal posteriors [17].

$$p(\mathbf{x}_k, \mathbf{r}_k|\mathbf{z}_{1:k}) \approx Q(\mathbf{x}_k, \mathbf{r}_k) = Q_x(\mathbf{x}_k)Q_r(\mathbf{r}_k) \quad (22)$$

where $Q_x(\mathbf{x}_k)$ and $Q_r(\mathbf{r}_k)$ are unknown approximating marginal densities of \mathbf{x}_k and \mathbf{r}_k .

The basic idea of VB approximation is to minimize the Kullback-Leibler (KL) divergence between the approximating posterior and the true posterior:

$$KL[Q_x(\mathbf{x}_k)Q_r(\mathbf{r}_k)||p(\mathbf{x}_k, \mathbf{r}_k|\mathbf{z}_{1:k})] = \int Q_x(\mathbf{x}_k)Q_r(\mathbf{r}_k) \times \log \left(\frac{Q_x(\mathbf{x}_k)Q_r(\mathbf{r}_k)}{p(\mathbf{x}_k, \mathbf{r}_k|\mathbf{z}_{1:k})} \right) d\mathbf{x}_k d\mathbf{r}_k \quad (23)$$

Given the measurements $\mathbf{z}_{1:k}$, we can minimize the KL divergence with respect to the probability densities $Q_x(\mathbf{x}_k)$ and $Q_r(\mathbf{r}_k)$ in turn, while keeping the other fixed. Then, the following equations can be given as:

$$Q_x(\mathbf{x}_k) \propto \exp(\langle \ln p(\mathbf{x}_k, \mathbf{r}_k, \mathbf{z}_k|\mathbf{z}_{1:k-1}) \rangle_{\mathbf{r}_k}) \quad (24)$$

$$Q_r(\mathbf{r}_k) \propto \exp(\langle \ln p(\mathbf{x}_k, \mathbf{r}_k, \mathbf{z}_k | \mathbf{z}_{1:k-1}) \rangle_{\mathbf{x}_k}) \quad (25)$$

where $\langle \cdot \rangle_{\mathbf{x}_k}$ and $\langle \cdot \rangle_{\mathbf{r}_k}$ denote the expectations with respect to $Q_x(\mathbf{x}_k)$ and $Q_r(\mathbf{r}_k)$, respectively. Obviously, it is not an explicit solution since the distribution of each parameter is dependent on the other and neither distributions is known. The mechanism of VB method is to firstly give the initial values of the parameters and then use expectation-maximum (EM) algorithm to iteratively calculate $Q_x(\mathbf{x}_k)$ and $Q_r(\mathbf{r}_k)$ until convergence. For the above CE models, $Q_x(\mathbf{x}_k)$ and $Q_r(\mathbf{r}_k)$ can be obtained from the following procedure [19]:

(1) The VB expectation step yields:

$$Q_x(\mathbf{x}_k) \propto f(\mathbf{x}_k, \mathbf{z}_k) e^{\langle \phi(\mathbf{r}_k) \rangle_{\mathbf{r}_k}^T u(\mathbf{x}_k, \mathbf{z}_k)} = p(\mathbf{x}_k | \mathbf{z}_k, \langle \phi(\mathbf{r}_k) \rangle_{\mathbf{r}_k}) \quad (26)$$

(2) The VB maximization step yields that $Q_r(\mathbf{r}_k)$ is conjugate and of the form

$$Q_r(\mathbf{r}_k) = h(\alpha_k, \beta_k^- g(\mathbf{r}_k))^{\beta_k} e^{\phi(\mathbf{r}_k)^T \alpha_k} \quad (27)$$

where, α_k and β_k are the hyper-parameters, and

$$\alpha_k = \alpha_k^- + \langle u(\mathbf{x}_k, \mathbf{z}_k) \rangle_{\mathbf{x}_k} \quad (28)$$

$$\beta_k = \beta_k^- + n \quad (29)$$

where n is the dimension of the measurement.

4. VB Based Adaptive Shifted Rayleigh Filter with Unknown Clutter Probability

Considering the system model (1) and the measurement model (8) described in Section 2, we adopt the VB method within the SRF framework to get the joint estimation of the target state and the clutter probability.

The core is to determine the posterior approximation $Q(\mathbf{x}_k, r_k, \xi)$. Assume factorization $Q(\mathbf{x}_k, r_k, \xi) \approx Q_x(\mathbf{x}_k)Q_r(r_k)Q_\xi(\xi_k)$, we can obtain $Q_x(\mathbf{x}_k)$, $Q_r(r_k)$ and $Q_\xi(\xi_k)$ at each time k through the following procedure.

(1) Optimization of $Q_x(\mathbf{x}_k)$ for fixed $Q(r_k, \xi)$.

First, by using the first form of the measurement likelihood (12), the complete-data likelihood is presented as

$$\begin{aligned} p(\mathbf{x}_k, z_k | r_k, \mathbf{z}_{1:k-1}) &= p(z_k | \mathbf{x}_k, r_k) p(\mathbf{x}_k | r_k, \mathbf{z}_{1:k-1}) \\ &= \left[\frac{1}{2\pi} r_k + f(\theta_k | \mathbf{x}_k)(1 - r_k) \right] \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \end{aligned} \quad (30)$$

Then according to (24), we can get

$$\begin{aligned} Q_x(\mathbf{x}_k) &\propto \exp\{\langle \ln p(z_k, \mathbf{x}_k | r_k, \mathbf{z}_{1:k-1}) \rangle_{r_k, \xi}\} \\ &= \left[\frac{1}{2\pi} \langle r_k \rangle_{r_k} + f(\theta_k | \mathbf{x}_k)(1 - \langle r_k \rangle_{r_k}) \right] \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ &= \frac{1}{2\pi} \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \langle r_k \rangle_{r_k} + f(\theta_k | \mathbf{x}_k) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})(1 - \langle r_k \rangle_{r_k}) \end{aligned} \quad (31)$$

$$\approx \frac{1}{2\pi} \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \langle r_k \rangle_{r_k} + \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}) f(\theta_k | \mathbf{z}_{1:k-1})(1 - \langle r_k \rangle_{r_k}) \quad (32)$$

The approximation sign in (32) is because the following:

$$\begin{aligned} f(\theta_k | \mathbf{x}_k) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) &= p(\theta_k | \mathbf{x}_k, \mathbf{z}_{1:k-1}) p(\mathbf{x}_k | \mathbf{z}_{1:k-1}) \\ &= p(\mathbf{x}_k | \theta_k, \mathbf{z}_{1:k-1}) p(\theta_k | \mathbf{z}_{1:k-1}) \\ &\approx \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}) f(\theta_k | \mathbf{z}_{1:k-1}) \end{aligned} \quad (33)$$

where $f(\theta_k|\mathbf{z}_{1:k-1})$ is derived in Appendix B.

Comparing (32) with the posterior density (14) of SRF, the difference lies in the weights. Except for a normalization constant, the clutter probability ξ used before in (14) has been replaced by $\langle r_k \rangle_{r_k}$ in (32), which is updated online.

(2) Optimization of $Q(r_k, \xi)$ for fixed $Q_x(\mathbf{x}_k)$.

We use the VB method again by factorizing $Q(r_k, \xi) \approx Q_r(r_k)Q_\xi(\xi)$. Assume the conjugate prior of r_k is binomial distributed with parameter ξ . That is,

$$p(r_k|\xi) = \xi^{r_k}(1 - \xi)^{1-r_k} \tag{34}$$

and $p(\xi)$ follows beta distribution with parameters α_1 and α_2 .

$$p(\xi; \alpha_1, \alpha_2) = \frac{1}{B(\alpha_1, \alpha_2)} \xi^{\alpha_1-1} (1 - \xi)^{\alpha_2-1} \tag{35}$$

where $B(\alpha_1, \alpha_2) = \Gamma(\alpha_1)\Gamma(\alpha_2)/\Gamma(\alpha_1 + \alpha_2)$.

To have the form of (19), the complete-data likelihood $p(\mathbf{x}_k, z_k|r_k, \mathbf{z}_{1:k-1})$ is re-derived using the second form (13) of $p(z_k|\mathbf{x}_k, r_k)$ as

$$\begin{aligned} p(\mathbf{x}_k, z_k|r_k, \mathbf{z}_{1:k-1}) &= (1/2\pi)^{r_k} [f(\theta_k|\mathbf{x}_k)]^{1-r_k} \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \\ &= f(\theta_k|\mathbf{x}_k) \mathcal{N}(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \exp\{r_k[\ln(1/2\pi) - \ln[f(\theta_k|\mathbf{x}_k)]]\} \end{aligned} \tag{36}$$

Rewriting the conjugate prior of r_k in the form of (20), we can get

$$\begin{aligned} p(r_k|\xi) &= \xi^{r_k}(1 - \xi)^{1-r_k} \\ &= (1 - \xi) \exp[r_k \ln(\frac{\xi}{1 - \xi})] \end{aligned} \tag{37}$$

Then applying (27), $Q_r(r_k)$ can be obtained as

$$Q_r(r_k) = (1 - \eta_k) \exp[r_k \ln(\frac{\eta_k}{1 - \eta_k})] \tag{38}$$

where, η_k is the hyper-parameter and updated as

$$\ln(\frac{\eta_k}{1 - \eta_k}) = \langle \ln(\frac{\xi}{1 - \xi}) \rangle_\xi + \langle [\ln(1/2\pi) - \ln[f(\theta_k|\mathbf{x}_k)]] \rangle_{\mathbf{x}_k} \tag{39}$$

Likewise, rewriting the conjugate prior of ξ in the form of (20), we can get

$$\begin{aligned} p(\xi) &= \frac{1}{B(\alpha_1, \alpha_2)} \xi^{\alpha_1-1} (1 - \xi)^{\alpha_2-1} \\ &= \frac{1}{B(\alpha_1, \alpha_2)} (1 - \xi)^{\alpha_1+\alpha_2-2} (1 - \xi)^{-(\alpha_1-1)} \xi^{\alpha_1-1} \\ &= \frac{1}{B(\alpha_1, \alpha_2)} (1 - \xi)^{\alpha_1+\alpha_2-2} \exp[(\alpha_1 - 1) \ln(\frac{\xi}{1 - \xi})] \end{aligned} \tag{40}$$

Then applying (27), $Q_\xi(\xi)$ can be obtained as

$$Q_\xi(\xi) = \frac{1}{B(\alpha'_1, \alpha'_2)} (1 - \xi)^{\alpha'_1+\alpha'_2-2} \exp[(\alpha'_1 - 1) \ln(\frac{\xi}{1 - \xi})] \tag{41}$$

with hyper-parameters

$$\alpha'_1 = \alpha_1 + \langle r_k \rangle_{r_k} \quad (42)$$

$$\alpha'_2 = \alpha_2 + n - \langle r_k \rangle_{r_k} \quad (43)$$

where n is the dimension of the measurement.

According to the approximated posteriors of $Q_r(r_k)$ and $Q_\xi(\xi)$, $\langle r_k \rangle_{r_k}$ and $\langle \ln(\frac{\xi}{1-\xi}) \rangle_\xi$ can be obtained as:

$$\langle r_k \rangle_{r_k} = \eta_k \quad (44)$$

$$\langle \ln(\frac{\xi}{1-\xi}) \rangle_\xi = \psi(\alpha'_1) - \psi(\alpha'_2) \quad (45)$$

where $\psi(\cdot)$ is the digamma function.

Taking expectation and covariance on the posterior $Q_x(\mathbf{x}_k)$, the conditional mean and covariance of the target state can then be obtained. We summarize the entire filtering procedure of the VB-based SRF (VB-SRF) in Algorithm 1.

Algorithm 1 : VB-SRF.

(1) Initialization: $\bar{\mathbf{x}}_{0|0}, \bar{\mathbf{P}}_{0|0}, \mathbf{Q}_v, \mathbf{Q}_w, \eta_0, \alpha_{1,0}, \alpha_{2,0}$

(2) Prediction:

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1} \bar{\mathbf{x}}_{k-1|k-1} + \mathbf{u}_{k-1}^s$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}_{k-1} \bar{\mathbf{P}}_{k-1|k-1} \mathbf{F}_{k-1}^T + \mathbf{Q}_v$$

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{Q}_k^m$$

$$\eta_{k|k-1} = \rho \eta_{k-1}, \quad \alpha_{1,k|k-1} = \rho \alpha_{1,k-1}, \quad \alpha_{2,k|k-1} = \rho \alpha_{2,k-1}$$

where ρ is the scale factor and $0 < \rho \leq 1$.

(3) Update: the update of VB-SRF utilizes iterate filtering framework.

(3.a) First set: $\bar{\mathbf{x}}_{k|k}^{(0)} = \hat{\mathbf{x}}_{k|k-1}, \bar{\mathbf{P}}_{k|k}^{(0)} = \mathbf{P}_{k|k-1}, \eta_k^{(0)} = \eta_{k|k-1}, \alpha_{1,k}^{(0)} = \alpha_{1,k|k-1}, \alpha_{2,k}^{(0)} = \alpha_{2,k|k-1}$

(3.b) Calculate state estimation and its covariance using SRF when the measurement is from the target:

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}$$

$$\varepsilon_k = (\mathbf{b}_k^T \mathbf{S}_k^{-1} \mathbf{b}_k)^{-1/2} \mathbf{b}_k^T \mathbf{S}_k^{-1} (\mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} + \mathbf{u}_k^m)$$

$$\gamma_k = (\mathbf{b}_k^T \mathbf{S}_k^{-1} \mathbf{b}_k)^{-1/2} \rho_n(\varepsilon_k)$$

$$\delta_k = (\mathbf{b}_k^T \mathbf{S}_k^{-1} \mathbf{b}_k)^{-1/2} [2 + \varepsilon_k \rho_2(\varepsilon_k) - \rho_2^2 \varepsilon_k]$$

$$\rho_2(\varepsilon_k) = \frac{\varepsilon_k e^{-\varepsilon_k^2/2} + \sqrt{2\pi}(\varepsilon_k^2 + 1) F_{normal}(\varepsilon_k)}{e^{-\varepsilon_k^2/2} + \sqrt{2\pi}(\varepsilon_k) F_{normal}(\varepsilon_k)}$$

$$\hat{\mathbf{x}}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \hat{\mathbf{x}}_{k|k-1} - \mathbf{K}_k \mathbf{u}_k^m + \gamma_k \mathbf{K}_k \mathbf{b}_k$$

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_{k|k-1} + \delta_k \mathbf{K}_k \mathbf{b}_k \mathbf{b}_k^T \mathbf{K}_k^T$$

(3.c) For $j = 1 : N$, iterate the following N (N denotes iterated times) steps:

• **Calculate the fused state estimation and its covariance:**

$$\bar{\mathbf{x}}_{k|k}^{(j)} = \frac{1}{2\pi c} \eta_k^{(j-1)} \hat{\mathbf{x}}_{k|k-1} + \frac{1}{c} (1 - \eta_k^{(j-1)}) f(\theta_k | \mathbf{z}_{1:k-1}) \hat{\mathbf{x}}_{k|k}$$

$$\bar{\mathbf{P}}_{k|k}^{(j)} = \frac{1}{2\pi c} \eta_k^{(j-1)} (\mathbf{P}_{k|k} + (\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}_{k|k}) (\hat{\mathbf{x}}_{k|k} - \bar{\mathbf{x}}_{k|k})^T) + \frac{1}{c} (1 - \eta_k^{(j-1)}) f(\theta_k | \mathbf{z}_{1:k-1}) (\mathbf{P}_{k|k-1} + (\hat{\mathbf{x}}_{k|k-1} - \bar{\mathbf{x}}_{k|k}^{(j)}) (\hat{\mathbf{x}}_{k|k-1} - \bar{\mathbf{x}}_{k|k}^{(j)})^T)$$

where $c = \frac{1}{2\pi} \eta_k^{(j-1)} + f(\theta_k) (1 - \eta_k^{(j-1)})$ is a normalization term, and $f(\theta_k | \mathbf{z}_{1:k-1})$ can be obtained using (A6).

• **Update parameters:**

$$\ln\left(\frac{\eta_k^{(j)}}{1 - \eta_k^{(j)}}\right) = \psi(\alpha_{1,k}^{(j-1)}) - \psi(\alpha_{2,k}^{(j-1)}) + \ln(1/2\pi) - \ln f(\theta_k | \bar{\mathbf{x}}_{k|k}^{(j)})$$

$$\alpha_{1,k}^{(j)} = \alpha_{1,k}^{(j-1)} + \eta_k^{(j)}$$

$$\alpha_{2,k}^{(j)} = \alpha_{2,k}^{(j-1)} - \eta_k^{(j)} + 1$$

• **End for and set** $\bar{\mathbf{x}}_{k|k} = \bar{\mathbf{x}}_{k|k}^{(N)}, \bar{\mathbf{P}}_{k|k} = \bar{\mathbf{P}}_{k|k}^{(N)}, \eta_k = \eta_k^{(N)}, \alpha_{1,k} = \alpha_{1,k}^{(N)}, \alpha_{2,k} = \alpha_{2,k}^{(N)}$.

5. Simulation Results

To evaluate the performance of the VB-SRF algorithm, two scenarios which are almost the same with these in [12,13] are utilized. The differences lie in the clutter probability in scenario 1 and the sensor tracks in scenario 2 which were not detailed in [12]. The two scenarios are very representative. In scenario 1, a maneuvering sensor is used in order to satisfy the condition of observability in bearings-only tracking. In scenario 2, multiple distributed sensors with large noise variance are utilized, which make the problem more challenging. The tracking performance of the VB-SRF algorithm was compared with the SRF algorithm, the MFPDA-SCKF algorithm and PDA-SCKF algorithm in terms of track loss, track accuracy and computation complexity. The track loss is declared when the track error is large enough that making the filter diverge. Root mean square (RMS) error is used to show the track accuracy. In addition, the computation complexity is reflected by the computation time of each filter. The simulation codes can be downloaded through Github [20].

5.1. Scenario 1

In scenario 1, a target moves along a horizontal track, with zero vertical displacement, according to a white noise acceleration model. The state of the target is represented as $\mathbf{x}_k = [x_{1,k}, \dot{x}_{1,k}]$, where $x_{1,k}$ and $\dot{x}_{1,k}$ are the horizontal distance and velocity at time k , respectively. The observer platform follows an approximately parallel track at a constant average speed. The horizontal and vertical displacements of the platform $x_k^p = [x_{1,k}^p, x_{2,k}^p]^T$ are governed by the equations:

$$x_{1,k}^p = 4k + \tilde{x}_{1,k}^p \quad (46)$$

$$x_{2,k}^p = 20 + \tilde{x}_{2,k}^p \quad (47)$$

in which $\tilde{x}_{1,k}^p$ and $\tilde{x}_{2,k}^p$ are zero mean Gaussian white noise processes, both with variance $q = 1$.

The measurement is the angle (in radians) of the line-of-sight of this target from the platform. The sensor noise is Gaussian white noise with variance $\sigma^2 = (0.05)^2 \text{rad}^2 = 2.86^2 \text{deg}^2$. The true clutter probability is 0.8. Other parameters are detailed in [13]. The configuration of the observer platform and target is illustrated in Figure 1. The bearing measurement of the target is presented in Figure 2. It can be seen that there is no abrupt change.

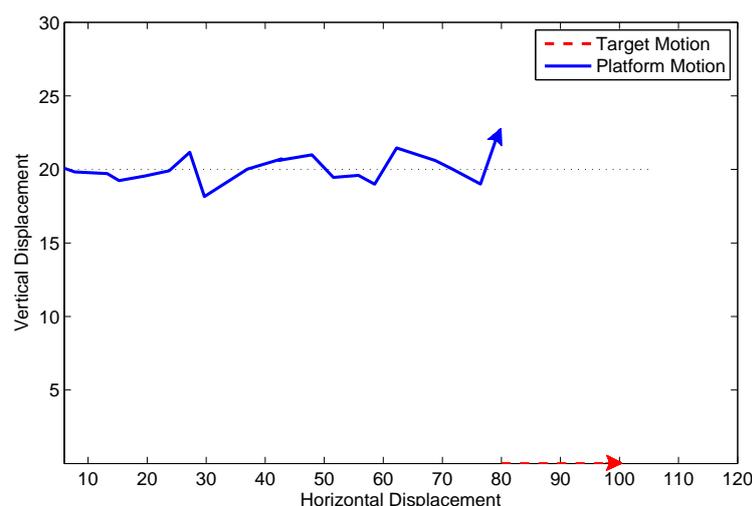


Figure 1. Target-observer geometry in Scenario 1.

The clutter probability used in SRF is set as $p_c = 0.8$, which is the same with the true clutter probability. The initial values of VB-SRF parameters are $\eta_0 = 0.8$, $\alpha_{1,0} = 2$, $\alpha_{2,0} = 10$. The clutter

density λ is calculated using $-k \ln(1 - p_c)/2\pi$ in MEFPDA-SCKF and PDA-SCKF. Figure 3 presents the RMS target position errors of the four filters using 1000 Monte Carlo runs. It can be seen that SRF and VB-SRF have comparable performance under the correct clutter probability. MEFPDA-SCKF and PDA-SCKF have a little better tracking accuracy than SRF and VB-SRF.

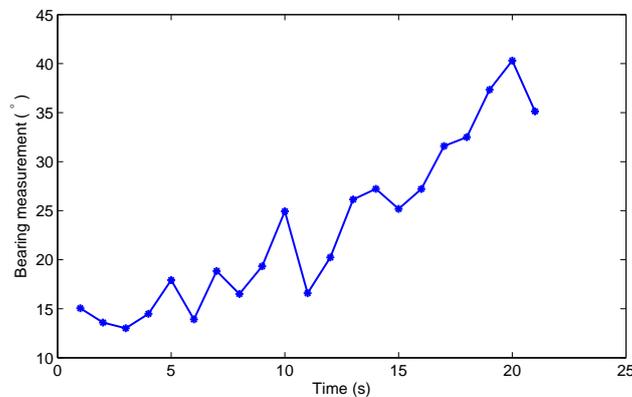


Figure 2. The measurement of the target in Scenario 1.

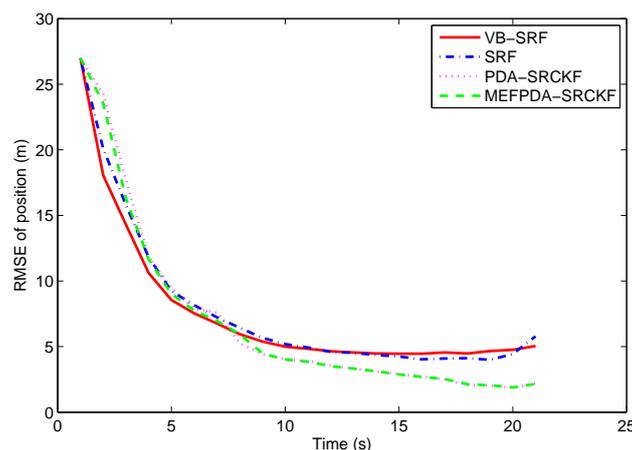


Figure 3. RMS target position errors with correct clutter probability in Scenario 1.

In challenging scenarios, the clutter probability maybe unknown or time varying. The pre-set parameters are probably inaccurate. So here we set mistuned clutter probabilities for the four filters: (1) $p_c = \eta_0 = 0.7$; (2) $p_c = \eta_0 = 0.5$; (3) $p_c = \eta_0 = 0.3$. The percentages of track losses in 1000 Monte Carlo runs are given in Table 1. Clearly, the VB-SRF algorithm outperforms the SRF algorithm due to fewer track losses. Moreover the proportion of track losses increases as the mistuning aggravates. Especially, there are 2.9% tracks are lost for SRF while only 0.1% tracks are lost for VB-SRF in the worst case. In addition, we can see that MEFPDA-SCKF and PDA-SCKF have no track loss in all the three mistuned cases.

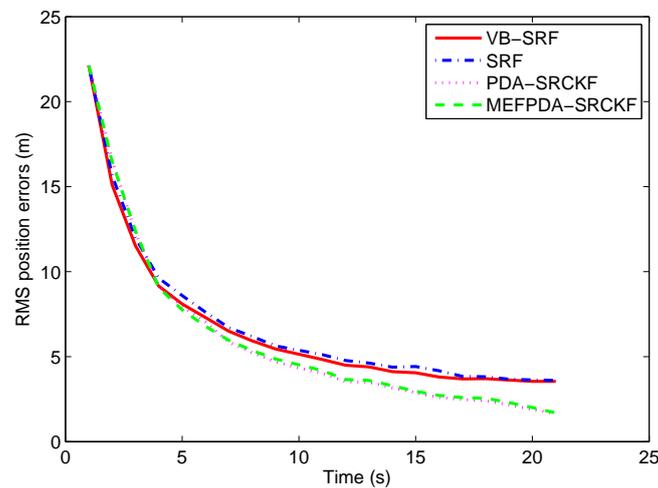
The RMS position errors of the four filters with different mistuned clutter probabilities are shown in Figure 4. We only consider the runs without track loss. From the figure, we can see that the tracking accuracy of VB-SRF is slightly better than SRF when $\eta_0 \geq 0.5$. However, the performance difference is not obvious. When $p_c = \eta_0 = 0.3$, VB-SRF exhibits distinct superiority over SRF. It implies that VB-SRF is more robust than SRF. Meanwhile, MEFPDA-SCKF and PDA-SCKF show better tracking accuracy than SRF and VB-SRF under all the three mistuned cases. They are almost not affected by the mistuning. It shows MEFPDA-SCKF and PDA-SCKF are more accurate and robust than SRF and VB-SRF under this simple scenario.

Table 2 shows the computation time of the four filters with 100 Monte Carlo runs. It is clear that VB-SRF has the maximum time of computation, which is about 2 times of that of SRF. The computation time of MEFPDA-SCKF and PDA-SCKF are comparable and both smaller than SRF and VB-SRF.

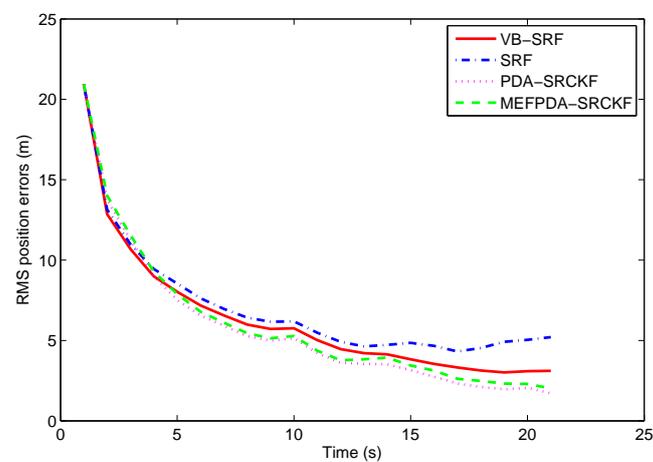
On the whole, the VB-SRF algorithm outperforms the SRF in terms of track continuity, track accuracy and robustness especially in severely mismatched scenarios but with higher computation complexity. MEFPDA-SCKF and PDA-SCKF perform better than SRF and VB-SRF in all aspects. It illustrates that the PDA-SCKF-based strategy has superiority over the SRF-based strategy in handling the clutters in simple scenarios.

Table 1. The percentages of track losses of four filters in two scenarios.

	Scenario 1			Scenario 2		
	$p_c = 0.7$	$p_c = 0.5$	$p_c = 0.3$	$p_c = 0.667$	$p_c = 0.5$	$p_c = 0.3$
VB-SRF	0	0	0.1%	0	0	0
SRF	0.9%	1.6%	2.9%	0	0	2.7%
MEFPDA-SCKF	0	0	0	13.5%	13.3%	14.2%
PDA-SCKF	0	0	0	20.1%	20.8%	22.4%

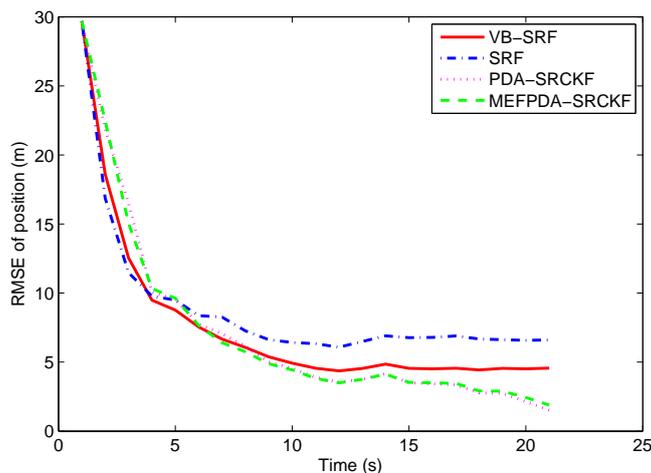


(a) $p_c = \eta_0 = 0.7$



(b) $p_c = \eta_0 = 0.5$

Figure 4. Cont.



(c) $p_c = \eta_0 = 0.3$

Figure 4. RMS target position errors with different mistuned clutter probabilities in Scenario 1.

Table 2. Computation time of the four filters with 100 Monte Carlo runs for two scenarios.

	Scenario 1	Scenario 2
VB-SRF	0.7406 s	1.0236 s
SRF	0.3690 s	0.5779 s
MEFPDA-SCKF	0.2066 s	0.3314 s
MEFPDA-SCKF	0.2092 s	0.3128 s

5.2. Scenario 2

For scenario 2, the aim is to track a single target from several drifting sonobuoys. A monitoring aircraft estimates the positions of the drifting sonobuoys by observing the direction of arrival of sensor transmissions. The sonobuoys track the position of the target by means of noisy bearings measurements. The state is 12-dimensional:

$$\mathbf{x}_k = [x_{0,k}, \dot{x}_{0,k}, y_{0,k}, \dot{y}_{0,k}, x_{1,k}, y_{1,k}, x_{2,k}, y_{2,k}, x_{3,k}, y_{3,k}, u_{1,k}, u_{1,k}]^T \tag{48}$$

the first four components represent the (x, y) coordinates of the position and velocity of the target, the next six, the coordinates of the positions of the three sonobuoys, and the last two, those of the drift current effecting all three sonobuoys.

Six simultaneous measurements are made at each time step. Three of these are measurements of the bearing angles of the sonobuoys from the monitoring platform and they are uncluttered. Three are the bearing angles of the target from the sonobuoys, which are subject to clutter. The standard deviation of monitoring sensor bearing noise and sonobuoy sensor bearing noise are 0.8° and 16° , respectively. The true probability of clutter is set as 0.667. In addition, the bearing of the clutter is uniformly distributed over $[-\pi, \pi]$. Other simulation parameters can be referred to [12]. 200 Monte Carlo runs are performed to evaluate the performance of the proposed filter. Figure 5 shows the behaviour of the estimates of target and sonobuoy positions provided by both the SRF and VB-SRF, for a typical simulation.

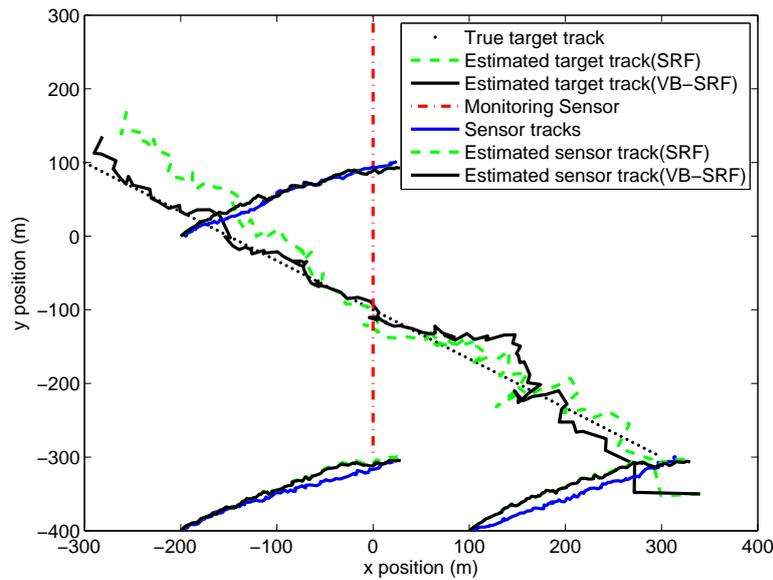


Figure 5. Typical tracks of target, drifting sonobuoy sensors, together with the estimated tracks.

The RMS target position errors of the four filters with correct clutter probability assumption are given in Figure 6. Compared with scenario 1, the differences between the four filters are more dramatic in scenario 2. This is probably because scenario 2 is more complex in which multiple sensors are used to observe the target and give abruptly changing and severely noise-corrupted bearing measurements of the target, shown in Figure 7. Thus, even a minor change in filtering strategy could result in large variations in performance. Meanwhile, seen from Figure 7, an abrupt change (almost from $+180^\circ$ to -180°) occurs in the target bearing measurement from sonobuoy sensor 3 at $k = 62$ s, which leads to several track losses shown in Table 1 and much larger position errors of MEFPDA-SCKF and PDA-SCKF. Whereas, SRF and VB-SRF are less affected by the abrupt bearing variation since the value of the projected measurement $b_k = (\sin\theta_k, \cos\theta_k)^T$ is invariant when there is a 360° change in bearing θ_k . They have comparable tracking accuracy. It is hard to decide which one is better.

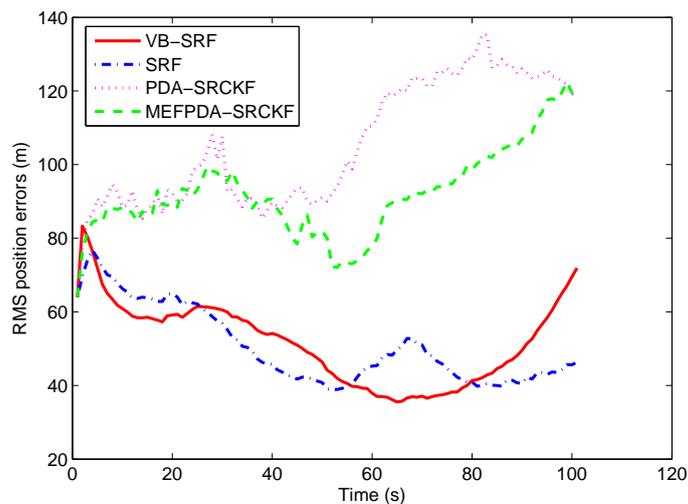


Figure 6. RMS target position errors with correct clutter probability in Scenario 2.

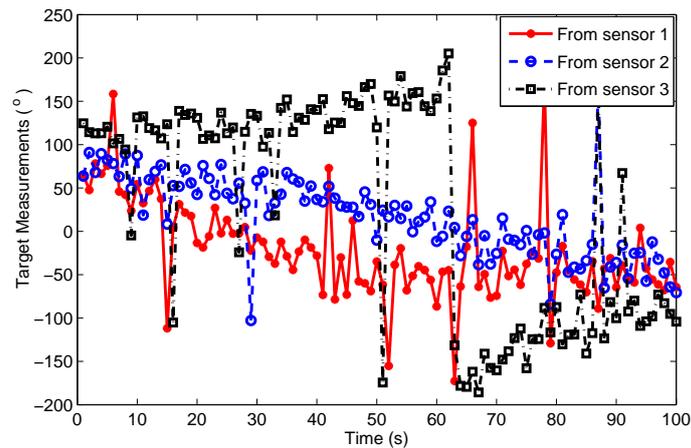
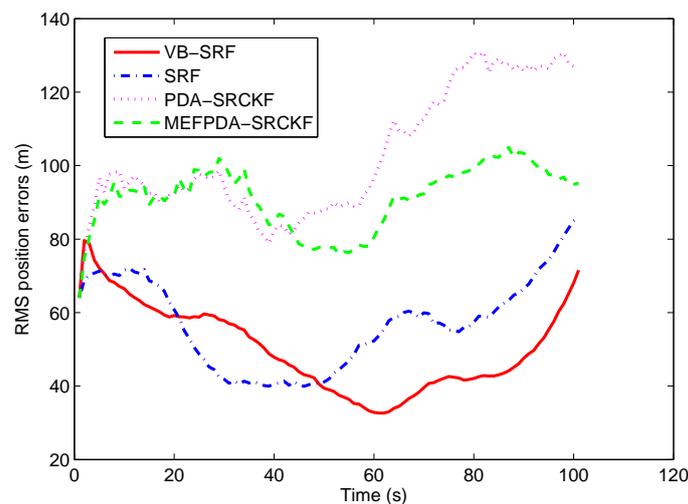


Figure 7. The target measurements from three sonobuoy sensors in Scenario 2.

To compare the performance under adverse scenarios, we set mistuned clutter probabilities as: (1) $p_c = \eta_0 = 0.5$; (2) $p_c = \eta_0 = 0.3$. In case (1), there is no track loss in SRF and VB-SRF and the RMS position errors of the two filters are shown in Figure 8a. We can see that the RMS position errors of SRF are slightly increased compared with the case with no mistuning, while the RMS position errors of VB-SRF remain almost unchanged. In case (2), as shown in Table 1, 2.7% tracks are lost for SRF while no track is lost for VB-SRF. Meanwhile, as can be seen in Figure 8b, VB-SRF has much smaller RMS errors than the SRF. For MEFPDA-SCKF and PDA-SCKF, unlike with scenario 1, they have higher percentage of track losses and larger RMS position errors than VB-SRF in both mistuned cases. It shows that VB-SRF is superior to PDA-SCKF-based algorithms in more challenging scenarios. In addition, from Table 2, we can see that the computation time of VB-SRF is twice of the SRF and three times of MEFPDA-SCKF and PDA-SCKF. Overall, all these reveal that the proposed VB-SRF algorithm has significant performance superiority in severely mismatched and complex cases at the cost of a little higher computation complexity.



(a) $p_c = \eta_0 = 0.5$

Figure 8. Cont.

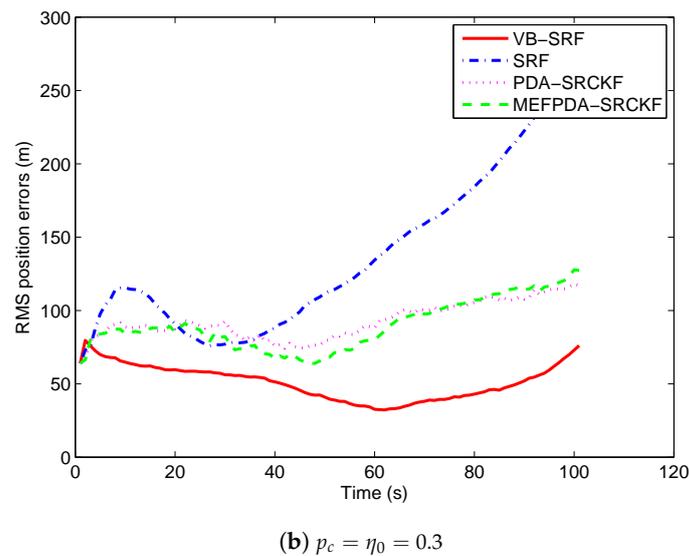


Figure 8. RMS target position errors with mistuned clutter probability in Scenario 2.

6. Conclusions

Bearings-only target tracking in the presence of clutter is a difficult problem because of the nonlinearity of the measurement model, the measurement origin uncertainty and the observability of the target. The Shifted Rayleigh filter (SRF) is shown to exhibit good performance for bearings-only target tracking in certain challenging scenarios through exploiting the essential structure of the nonlinearities in a new way. However, the clutter probability is assumed known and constant in SRF, which may not match with the truth especially in adverse scenarios. Therefore, to handle the bearings-only target tracking in clutters with uncertain clutter probability, a variational Bayesian-based adaptive shifted Rayleigh filter (VB-SRF) is proposed in this paper. By establishing a conjugate exponential model of the clutter probability and the data association indicator, the approximated posterior probability densities of the target state and the clutter parameters are iteratively calculated using the VB expectation and maximization steps. Finally, joint estimation of the target state and the clutter probability are achieved in the framework of SRF. The tracking performance of the proposed filter is compared with SRF, PDA-SCKF and MEFPDA-SCKF via two simulation examples. It shows that the proposed filter outperforms the other three filters in terms of track continuity and track accuracy with a little higher computation complexity in complex adverse scenarios. In addition, it also reveals that the proposed VB-SRF exhibits better robustness than the traditional SRF.

Author Contributions: J.H. conceived this paper, derived the method, and wrote the original draft ; Y.Y. made the investigation; and T.G. revised the paper and provided some valuable suggestions.

Funding: This research was funded by Aeronautical Science Foundation of China (No. 2016ZC53033), and Shaanxi Provincial Key Research and Development Programs (No. 2017ZDXM-GY-06, No. 2017GY-057), and Xi'an Science and Technology Planning Project—Scientific and Technological Innovation Guidance Project (No. 201805042YD20CG26 (8)).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

VB	Variational Bayesian
SRF	Shifted Rayleigh Filter
PDA	Probability Data Association
EKF	Extended Kalman Filter

MPEKF	Polar Coordinate EKF
PLE	Pseudo-Linear Estimator
UKF	Unscented Kalman Filter
CKF	Cubature Kalman Filter
PF	Particle Filter
MEFPDA	Maximum Entropy Fuzzy Probabilistic Data Association
SCKF	Square-root Cubature Kalman Filter
CE	Conjugate Exponential
KL	Kullback- Leibler
EM	Expectation-Maximum
RMS	Root Mean Square

Appendix A. Derivation of $f(\theta_k | \mathbf{x}_k)$

Given the true state \mathbf{x}_k , the “augmented” measurement \mathbf{y}_k is assumed to be $\mathcal{N}(\mathbf{m}_k, \mathbf{Q}_w)$ variable, where $\mathbf{m}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{u}_k^m$. θ_k is the bearing of \mathbf{y}_k .

Let $\mathbf{y}'_k = \mathbf{Q}_w^{-1/2} \mathbf{y}_k$, then $\mathbf{y}'_k \sim \mathcal{N}(\mathbf{Q}_w^{-1/2} \mathbf{m}_k, \mathbf{I})$. Its bearing has a density about its mean of the form $\alpha_\sigma(\theta)$. More precisely, let

$$g_k(\theta) = [(\sin\theta, \cos\theta) \mathbf{Q}_w^{-1} (\sin\theta, \cos\theta)^T]^{-1/2} \quad (\text{A1})$$

and define

$$h_k(\theta) = \arctan \left(\frac{a_{11} \sin\theta + a_{12} \cos\theta}{a_{21} \sin\theta + a_{22} \cos\theta} \right) \quad (\text{A2})$$

where $[a_{ij}] = \mathbf{Q}_w^{-1/2}$.

Actually, $g_k(\theta)$ and $h_k(\theta)$ are the reciprocal length and bearing of the transformed unit vector $\frac{\mathbf{Q}_w^{-1/2} \mathbf{y}_k}{\|\mathbf{y}_k\|}$. The bearing $h_k(\theta_k)$ then has $\alpha_{g_k(\theta_k^m)}(h_k(\theta_k) - h_k(\theta_k^m))$ as its density, where θ_k^m is the bearing of \mathbf{m}_k . Inserting a Jacobian term, we can obtain the likelihood of measurement θ_k :

$$f(\theta_k | \mathbf{x}_k) = \frac{g_k^2(\theta_k)}{(\det \mathbf{Q}_w)^{1/2}} \alpha_{g_k(\theta_k^m)}(h_k(\theta_k) - h_k(\theta_k^m)) \quad (\text{A3})$$

Appendix B. Derivation of $f(\theta_k | \mathbf{z}_{1:k-1})$

Given the previous measurements $\mathbf{y}_{1:k-1}$, the “augmented” measurement \mathbf{y}_k is assumed to be $\mathcal{N}(\hat{\mathbf{y}}_k, \mathbf{S}_k)$ variable, where $\hat{\mathbf{y}}_k = \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1} + \mathbf{u}_k^m$, $\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_{k|k-1} \mathbf{H}_k^T + \mathbf{Q}_w$.

Let $\mathbf{y}'_k = \mathbf{S}_k^{-1/2} \mathbf{y}_k$, then $\mathbf{y}'_k \sim \mathcal{N}(\mathbf{S}_k^{-1/2} \hat{\mathbf{y}}_k, \mathbf{I})$. The bearing of \mathbf{y}'_k has a density about its mean of the form $\alpha_\sigma(\theta)$.

Let

$$g'_k(\theta) = [(\sin\theta, \cos\theta) \mathbf{S}_k^{-1} (\sin\theta, \cos\theta)^T]^{-1/2} \quad (\text{A4})$$

$$h'_k(\theta) = \arctan \left(\frac{s_{11} \sin\theta + s_{12} \cos\theta}{s_{21} \sin\theta + s_{22} \cos\theta} \right) \quad (\text{A5})$$

where $[s_{ij}] = \mathbf{S}_k^{-1/2}$.

$g'_k(\theta)$ and $h'_k(\theta)$ are the reciprocal length and bearing of the transformed unit vector $\frac{\mathbf{S}_k^{-1/2} \mathbf{y}_k}{\|\mathbf{y}_k\|}$. The bearing $h'_k(\theta_k)$ then has $\alpha_{g'_k(\hat{\theta}_k)}(h'_k(\theta_k) - h'_k(\hat{\theta}_k))$ as its density, where $\hat{\theta}_k$ is the bearing of $\hat{\mathbf{y}}_k$. Inserting a Jacobian term, we can obtain the probability density function of measurement θ_k given previous measurements $\mathbf{z}_{1:k-1}$:

$$f(\theta_k | \mathbf{z}_{1:k-1}) = \frac{g_k'^2(\theta_k)}{(\det \mathbf{S}_k)^{1/2}} \alpha_{g'_k(\hat{\theta}_k)}(h'_k(\theta_k) - h'_k(\hat{\theta}_k)) \quad (\text{A6})$$

References

1. Leong, P.H.; Arulampalam, S.; Lamahewa, T.A.; Abhayapala, T.D. A Gaussian-sum based cubature Kalman filter for bearings-only tracking. *IEEE Trans. Aerosp. Electron. Syst.* **2013**, *49*, 1161–1176. [CrossRef]
2. Aidala, V.J. Kalman filter behavior in bearings-only tracking applications. *IEEE Trans. Aerosp. Electron. Syst.* **1979**, *15*, 29–39. [CrossRef]
3. Aidala, V.J.; Hammel, S. Utilization of modified polar coordinates for bearings-only tracking. *IEEE Trans. Autom. Control* **1983**, *28*, 283–294. [CrossRef]
4. Aidala, V.J.; Nardone, S.C. Biased estimation properties of the pseudo linear tracking filter. *IEEE Trans. Aerosp. Electron. Syst.* **1982**, *18*, 432–441. [CrossRef]
5. Doğançay, K. On the efficiency of a bearings-only instrumental variable estimator for target motion analysis. *Signal Process.* **2005**, *85*, 481–490. [CrossRef]
6. Doğançay, K. Bias compensation for the bearings-only pseudolinear target track estimator. *IEEE Trans. Signal Process.* **2006**, *54*, 59–68. [CrossRef]
7. Nguyen, N.H.; Doğançay, K. Improved pseudolinear Kalman filter algorithms for bearings-only target tracking. *IEEE Trans. Signal Process.* **2017**, *65*, 6119–6134. [CrossRef]
8. Wang, W.P.; Liao, S.; Xing, T.W. The unscented Kalman filter for state estimation of 3-dimension bearing-only tracking. In Proceedings of the 2009 International Conference on Information Engineering and Computer Science, Wuhan, China, 19–20 December 2009; pp. 1–5.
9. Yang, R.; Ng, G.W.; Bar-Shalom, Y. Bearings-only tracking with fusion from heterogenous passive sensors: ESM/EO and acoustic. In Proceedings of the 2015 18th International Conference on Information Fusion (Fusion), Washington, DC, USA, 6–9 July 2015; pp. 1810–1816.
10. Hong, S.H.; Shi, Z.G.; Chen, K.S. Novel roughening algorithm and hardware architecture for bearings-only tracking using particle filter. *J. Electromagn. Waves Appl.* **2008**, *22*, 411–422. [CrossRef]
11. Chang, D.C.; Fang, M.W. Bearing-only maneuvering mobile tracking with nonlinear filtering algorithms in wireless sensor networks. *IEEE Syst. J.* **2014**, *8*, 160–170. [CrossRef]
12. Clark, J.M.C.; Vinter, R.B.; Yaqoob, M.M. The shifted Rayleigh filter for bearings only tracking. In Proceedings of the 2005 7th International Conference on Information Fusion, Philadelphia, PA, USA, 25–28 July 2005; pp. 93–100.
13. Clark, J.M.C.; Vinter, R.B.; Yaqoob, M.M. Shifted Rayleigh filter: A new algorithm for bearings-only tracking. *IEEE Trans. Aerosp. Electron. Syst.* **2007**, *43*, 1373–1384. [CrossRef]
14. Arulampalam, S.; Clark, M.; Vinter, R. Performance of the shifted Rayleigh filter in single-sensor bearings-only tracking. In Proceedings of the 2007 10th International Conference on Information Fusion, Quebec, QC, Canada, 9–12 July 2007; pp. 1–6.
15. Mei, D.; Liu, K.; Wang, Y. MEFPDA-SCKF for underwater single observer bearings-only target tracking in clutter. In Proceedings of the 2013 OCEANS—San Diego, San Diego, CA, USA, 23–27 September 2013; pp. 1–6.
16. Sarkka, S.; Nummenmaa, A. Recursive noise adaptive Kalman filtering by variational Bayesian approximations. *IEEE Trans. Autom. Control* **2009**, *54*, 596–600. [CrossRef]
17. Li, K.; Chang, L.; Hu, B. A variational Bayesian-based unscented Kalman filter with both adaptivity and robustness. *IEEE Sens. J.* **2016**, *16*, 6966–6976. [CrossRef]
18. Sun, J.; Zhou, J.; Li, X.R. State estimation for systems with unknown inputs based on variational Bayes method. In Proceedings of the 2012 15th International Conference on Information Fusion, Singapore, 9–12 July 2012; pp. 983–990.
19. Beal, M.J. Variational Algorithms for Approximate Bayesian Inference. Ph.D. Dissertation, Gatsby Computational Neuroscience Unit, University College London, London, UK, 2003.
20. Hou, J. jaylin254/VB-SRF1. Available online: <https://github.com/jaylin254/VB-SRF1> (accessed on 23 March 2019).

