

Article

Adaptive Maximum Correntropy Gaussian Filter Based on Variational Bayes

Guoqing Wang ¹, Zhongxing Gao ^{1,*}, Yonggang Zhang ¹ and Bin Ma ²

¹ College of Automation, Harbin Engineering University, Harbin 150001, China; wangguoqing2014@hrbeu.edu.cn (G.W.); zhangyg@hrbeu.edu.cn (Y.Z.)

² Chinese Ship Research and Design Center, Wuhan 430064, China; masanshao@163.com

* Correspondence: zhongxing1141@sina.com

Received: 20 May 2018; Accepted: 14 June 2018; Published: 17 June 2018



Abstract: In this paper, we investigate the state estimation of systems with unknown covariance non-Gaussian measurement noise. A novel improved Gaussian filter (GF) is proposed, where the maximum correntropy criterion (MCC) is used to suppress the pollution of non-Gaussian measurement noise and its covariance is online estimated through the variational Bayes (VB) approximation. MCC and VB are integrated through the fixed-point iteration to modify the estimated measurement noise covariance. As a general framework, the proposed algorithm is applicable to both linear and nonlinear systems with different rules being used to calculate the Gaussian integrals. Experimental results show that the proposed algorithm has better estimation accuracy than related robust and adaptive algorithms through a target tracking simulation example and the field test of an INS/DVL integrated navigation system.

Keywords: Gaussian filter; maximum correntropy criterion; variational Bayes; Kalman filter

1. Introduction

As the benchmark work in state estimation problems, the linear recursive Kalman filter (KF) has been applied in various applications, such as information fusion, system control, integrated navigation, target tracking, and GPS solutions [1–4]. It is then extended to nonlinear systems through different ways to approximate the nonlinear functions or filtering distributions. Using the Taylor series to linearize the nonlinear functions, the popular extended Kalman filter (EKF) [5] is obtained. To further improve the estimation accuracy of EKF, several sigma points based nonlinear filters have been proposed in recent decades, such as unscented Kalman filter (UKF) using unscented transform [6], cubature Kalman filter (CKF) according to cubature rules [7], and divided difference Kalman filter (DDKF) adopting the polynomial approximations [8]. All these filters can be regarded as special cases of Gaussian filter (GF) [9–11], where the noise distribution is assumed to be Gaussian.

However, when the measurements are polluted by non-Gaussian noise, such as impulsive inference or outliers, GF will have worse estimation results and even break down [12,13]. Besides the computation extensive methods including particle filter [14,15], Gaussian sum filter [16], and multiple model filters [17], the robust filters, such as Huber's KF (HKF, also known as M-estimation) [18–20] and H_∞ filter [21], are also intended for the contaminated measurements. Although the H_∞ filter can obtain guaranteed bounded estimation error, it does not perform well under Gaussian noise [22]. The Huber's M-estimation is a combined l_1 and l_2 norm filter that can effectively suppress the non-Gaussian noise [18–20]. Recently, the information theoretical measure correntropy has been used to incorporate the non-Gaussian noise [23–28]. According to the maximum correntropy criterion (MCC), a new robust filter known as the maximum correntropy Kalman filter (MCKF) has been proposed in [24], and it is also extended to nonlinear systems using EKF [27] and UKF [25,26,28]. Simulation results

show that an MCC based GF (MCGF) may obtain better estimation accuracy than M-estimation when choosing a proper kernel bandwidth [23–26]. Even so, both MCGF and M-estimation still require the information of nominal measurement noise covariance, which may be unknown or time varying in some applications. In this situation, the performance of MCGF will degrade as shown in our simulation examples.

Traditionally, the unknown noise statistic can be estimated by the adaptive filter, such as the Sage–Husa filter [29] and fading memory filter [30]. One drawback of these recursion adaptive filters is that the previously estimated statistic of the last time instant will influence current estimation, which is not suitable for the case measurement noise having frequently changing statistics [31]. The recently proposed variational Bayes (VB) based adaptive filter avoids this limitation by VB approximation, and VB based GFs (VBGFs) for both linear [32] and nonlinear systems [9,33] have been proposed.

In this paper, we proposed an adaptive MCGF based on the VB approximation, which is especially useful for estimating the system state from the measurements with unknown covariance non-Gaussian noise. Typical applications include low-cost INS/GPS integrated navigation systems [32] and maneuvering target tracking [33]. To overcome the limitation of MCGF under unknown time varying measurement noise covariance, the VB method is utilised to improve the adaptivity of MCGF, which is achieved through the fixed-point iteration framework. As will be demonstrated in our simulation results, our proposed method has better estimation accuracy than related algorithms. Furthermore, various filters can be obtained by using different ways to calculate the Gaussian integrals.

The rest of this paper is given as follows. In Section 2, after briefly introducing the concept of correntropy, we give the general MCGF algorithm. In Section 3, we explain the main idea of the VB method and the procedure of embedding it into MCGF to obtain our proposed adaptive MCGF. Section 4 gives the experimental results of a typical target tracking model and an INS/DVL navigation system comparing with several related algorithms. Conclusions are made in the final section.

2. Gaussian Filter Based on the Maximum Correntropy Criterion

2.1. Correntropy

As a kind of similarity measure, the correntropy of random variables X and Y is defined as [24–26]

$$V(X, Y) = E[\kappa(X, Y)] = \int \int \kappa(x, y) dF_{X, Y}(x, y), \quad (1)$$

where $E[\cdot]$ denotes expectation, $F_{X, Y}(x, y)$ is the joint density function, and $\kappa(x, y)$ represents the Mercer kernel. The most popular Gaussian kernel is given as the following:

$$\kappa(x, y) = G_{\sigma}(e) = \exp\left(-\frac{e^2}{2\sigma^2}\right), \quad (2)$$

where $e = x - y$, and $\sigma > 0$ is the kernel bandwidth.

Then, taking the Taylor series expansion on the Gaussian correntropy, we obtain that

$$V(X, Y) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2^n \sigma^{2n} n!} E[(X - Y)^{2n}]. \quad (3)$$

Obviously, it contains all the even moments of $X - Y$ weighted by the kernel bandwidth σ . It enables us to capture high order information when applying the correntropy in signal processing. In practice, we can use the sampling data to estimate the real correntropy since the joint density function is usually unavailable.

2.2. Maximum Correntropy Gaussian Filter

In this paper, we consider the following nonlinear system with additive noise

$$\begin{aligned} x_i &= f_{i-1}(x_{i-1}) + w_{i-1}, \\ z_i &= h_i(x_i) + v_i, \end{aligned} \tag{4}$$

where $x_i \in \mathbb{R}^n$ is the system state at time i and $z_i \in \mathbb{R}^m$ denotes the measurement. $f_i(\cdot)$ and $h_i(\cdot)$ represent the known nonlinear functions. In standard GF, the process noise w_i and the measurement noise v_i are assumed to be zero mean Gaussian noise sequences with known covariance Q_i and R_i . The initial state x_0 has known mean \bar{x}_0 and covariance P_0 .

In both GF and MCGF, the one step estimation \hat{x}_i^- and its estimation covariance P_i^- are obtained through:

$$\hat{x}_i^- = \int f_{i-1}(x_{i-1}) \mathcal{N}(x_{i-1} | \hat{x}_{i-1}, P_{i-1}) dx_{i-1}, \tag{5}$$

$$P_i^- = \int (f_{i-1}(x_{i-1}) - \hat{x}_i^-)(f_{i-1}(x_{i-1}) - \hat{x}_i^-)^T \mathcal{N}(x_{i-1} | \hat{x}_{i-1}, P_{i-1}) dx_{i-1} + Q_{i-1}. \tag{6}$$

To further improve the robustness of GF, MCC has been applied on the derivation of measurement update of MCGF. Consider the following regression model based on Equations (4)–(6) [25]:

$$\begin{bmatrix} \hat{x}_i^- \\ z_i \end{bmatrix} = \begin{bmatrix} x_i \\ h_i(x_i) \end{bmatrix} + \begin{bmatrix} \delta x_i \\ v_i \end{bmatrix} \tag{7}$$

where $\delta x_i = \hat{x}_i^- - x_i$. The covariance of $[\delta x_i \ v_i]^T$ is

$$\begin{bmatrix} P_i^- & 0 \\ 0 & R_i \end{bmatrix} = \begin{bmatrix} S_{p,i} S_{p,i}^T & 0 \\ 0 & S_{r,i} S_{r,i}^T \end{bmatrix} = S_i S_i^T. \tag{8}$$

Multiplying S_i^{-1} on both sides of Equation (7), we obtain

$$e_i = D_i - W_i, \tag{9}$$

where $D_i = S_i^{-1} \begin{bmatrix} \hat{x}_i^- \\ z_i \end{bmatrix}$ and $W_i = S_i^{-1} \begin{bmatrix} x_i \\ h_i(x_i) \end{bmatrix}$.

Then, the optimal estimation \hat{x}_i under the MCC can be obtained through the following optimization problem:

$$\hat{x}_i = \arg \max_{x_i} J_L = \arg \max_{x_i} \sum_{k=1}^{m+n} G_\sigma(e_i(k)), \tag{10}$$

where $e_i(k)$ is the k -th element of e_i .

Equation (10) can be solved by:

$$\frac{\partial J_L}{\partial x_i} = \left(\frac{\partial h_i(x_i)}{\partial x_i} \right)^T C_i (D_i - W_i) = 0, \tag{11}$$

where

$$C_i = \begin{bmatrix} \tilde{C}_{x,i} & \mathbf{0} \\ \mathbf{0} & \tilde{C}_{y,i} \end{bmatrix}, \tag{12}$$

$$\tilde{C}_{x,i} = \text{diag}(G_\sigma(e_i(1)), \dots, G_\sigma(e_i(n))), \tag{13}$$

$$\tilde{C}_{y,i} = \text{diag}(G_\sigma(e_i(n+1)), \dots, G_\sigma(e_i(n+m))). \tag{14}$$

Note here that the $\text{diag}(\cdot)$ is used to denote the diagonal matrix.

Based on the above equation and using \hat{x}_i^- to replace the x_i contained in Equation (9), MCGF can be written in a similar way as GF except a modified measurement noise covariance [25]:

$$\tilde{\mathbf{R}}_i = \mathbf{S}_{r,i} \tilde{\mathbf{C}}_{y,i}^{-1} \mathbf{S}_{r,i}^T. \quad (15)$$

Then, \hat{x}_i and its covariance \mathbf{P}_i can be obtained through

$$\hat{x}_i = \hat{x}_i^- + \mathbf{K}_i(z_i - \hat{z}_i), \quad (16)$$

$$\mathbf{P}_i = \mathbf{P}_i^- - \mathbf{K}_i \mathbf{P}_{zz} \mathbf{K}_i^T, \quad (17)$$

where

$$\mathbf{K}_i = \mathbf{P}_{xz} \mathbf{P}_{zz}^{-1}, \quad (18)$$

$$\hat{z}_i = \int h_i(x_i) \mathcal{N}(x_i | \hat{x}_i^-, \mathbf{P}_i^-) dx_i, \quad (19)$$

$$\mathbf{P}_{zz} = \int (h_i(x_i) - \hat{z}_i)(h_i(x_i) - \hat{z}_i)^T \mathcal{N}(x_i | \hat{x}_i^-, \mathbf{P}_i^-) dx_i + \tilde{\mathbf{R}}_i, \quad (20)$$

$$\mathbf{P}_{xz} = \int (x_i - \hat{x}_i^-)(h_i(x_i) - \hat{z}_i)^T \mathcal{N}(x_i | \hat{x}_i^-, \mathbf{P}_i^-) dx_i. \quad (21)$$

We easily find that the main difference between MCGF and GF is the modified measurement noise covariance, and MCGF shows excellent estimation performance when measurement is polluted by outliers or shot noise [24–26]. However, it still requires the knowledge of measurement noise covariance. When the covariance changes over time (which implies the true covariance is different from the known covariance), the MCGF algorithm does not perform well. Therefore, we adopt the adaptive method to further improve the performance of MCGF in this case.

3. Variation Bayesian Maximum Correntropy Gaussian Filter

The main idea under state estimation is to obtain the posterior probability density function $p(x_i | z_{1:i})$. For GF, we obtain it through the Gaussian approximation $p(x_i | z_{1:i}) \approx \mathcal{N}(x_i | \hat{x}_i, \mathbf{P}_i)$. However, if the measurement noise covariance \mathbf{R}_i is unavailable, we need to estimate the joint posterior distribution $p(x_i, \mathbf{R}_i | z_{1:i})$. This distribution can be solved by the free form VB approximation [9,32]:

$$p(x_i, \mathbf{R}_i | z_{1:i}) \approx Q(x_i)Q(\mathbf{R}_i), \quad (22)$$

where $Q(x_i)$ and $Q(\mathbf{R}_i)$ are unknown approximation densities, which can be calculated by minimizing the Kullback–Leibler (KL) divergence between the true one and its corresponding approximation [9,32]:

$$Q(x_i) \propto \exp \left(\int \log p(z_i, x_i, \mathbf{R}_i | z_{1:i-1}) Q(\mathbf{R}_i) d\mathbf{R}_i \right), \quad (23)$$

$$Q(\mathbf{R}_i) \propto \exp \left(\int \log p(z_i, x_i, \mathbf{R}_i | z_{1:i-1}) Q(x_i) dx_i \right). \quad (24)$$

According to the VB method, $p(x_i, \mathbf{R}_i | z_{1:i})$ can be approximated as a product of Gaussian distribution and inverse Wishart (IW) distribution [9,32]:

$$p(x_i, \mathbf{R}_i | z_{1:i}) \approx \mathcal{N}(x_i | \hat{x}_i, \mathbf{P}_i) \text{IW}(\mathbf{R}_i | v_i, \mathbf{V}_i), \quad (25)$$

where

$$\mathcal{N}(x_i | \hat{x}_i, \mathbf{P}_i) \propto |\mathbf{P}_i|^{-1/2} \exp \left(-\frac{1}{2} (x_i - \hat{x}_i)^T \mathbf{P}_i^{-1} (x_i - \hat{x}_i) \right) \quad (26)$$

$$\text{IW}(\mathbf{R}_i | v_i, \mathbf{V}_i) \propto |\mathbf{R}_i|^{-(v_i+n+1)/2} \exp \left(-\frac{1}{2} \text{tr}(\mathbf{V}_i \mathbf{R}_i) \right), \quad (27)$$

where $\text{tr}(\cdot)$ is the trace of a matrix, and v_i and V_i are the degree of freedom parameter and the inverse scale matrix, respectively.

The integrals in (22) and (23) can be computed as follows [9,32]:

$$\begin{aligned} & \int \log p(\mathbf{z}_i, \mathbf{x}_i, \mathbf{R}_i | \mathbf{z}_{1:i-1}) Q(\mathbf{R}_i) d\mathbf{R}_i \\ &= -\frac{1}{2}(\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^T \langle \mathbf{R}_i^{-1} \rangle_R (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)) - \frac{1}{2}(\mathbf{x}_i - \hat{\mathbf{x}}_i^-)^T (\mathbf{P}_i^-)^{-1} (\mathbf{x}_i - \hat{\mathbf{x}}_i^-) + C_1, \end{aligned} \quad (28)$$

$$\begin{aligned} & \int \log p(\mathbf{z}_i, \mathbf{x}_i, \mathbf{R}_i | \mathbf{z}_{1:i-1}) Q(\mathbf{x}_i) d\mathbf{x}_i \\ &= -\frac{1}{2}(v_i^- + n + 2) \log |\mathbf{R}_i| - \frac{1}{2} \text{tr}(\mathbf{V}_i^- \mathbf{R}_i^{-1}) - \frac{1}{2} \langle (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)) \rangle_x + C_2, \end{aligned} \quad (29)$$

where $\langle \cdot \rangle_R = \int (\cdot) Q(\mathbf{R}_i) d\mathbf{R}_i$, $\langle \cdot \rangle_x = \int (\cdot) Q(\mathbf{x}_i) d\mathbf{x}_i$, and C_1, C_2 are some constants. Due to the fact that $Q(\mathbf{R}_i) = \text{IW}(\mathbf{R}_i | v_i, \mathbf{V}_i)$, we obtain

$$\langle \mathbf{R}_i^{-1} \rangle_R = (v_i - n - 1) \mathbf{V}_i^{-1}. \quad (30)$$

Besides this, the expectation can be rewritten as

$$\langle (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1} (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)) \rangle_x = \text{tr} \{ \langle (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^T (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)) \rangle_x \mathbf{R}_i^{-1} \}. \quad (31)$$

Substituting (30) and (31) into (28) and (29), and matching the parameters in (26) and (27), we can obtain the following results:

$$v_i = v_i^- + 1, \quad (32)$$

$$\mathbf{T}_i = \int (\mathbf{h}_i(\mathbf{x}_i) - \hat{\mathbf{z}}_i)(\mathbf{h}_i(\mathbf{x}_i) - \hat{\mathbf{z}}_i)^T \mathcal{N}(\mathbf{x}_i | \hat{\mathbf{x}}_i^-, \mathbf{P}_i^-) d\mathbf{x}_i + (v_i - n - 1)^{-1} \mathbf{V}_i, \quad (33)$$

$$\mathbf{K}_i = \mathbf{P}_{xz} (\mathbf{T}_i)^{-1}, \quad (34)$$

$$\hat{\mathbf{x}}_i = \hat{\mathbf{x}}_i^- + \mathbf{K}_i (\mathbf{z}_i - \hat{\mathbf{z}}_i), \quad (35)$$

$$\mathbf{P}_i = \mathbf{P}_i^- - \mathbf{K}_i \mathbf{T}_i \mathbf{K}_i^T, \quad (36)$$

$$\mathbf{V}_i = \mathbf{V}_i^- + \int (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)) (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^T \mathcal{N}(\mathbf{x}_i | \hat{\mathbf{x}}_i^-, \mathbf{P}_i^-) d\mathbf{x}_i, \quad (37)$$

where $(v_i - n - 1)^{-1} \mathbf{V}_i$ is the estimated measurement covariance.

The VB based GF works well for unknown measurement noise covariance. However, when the measurement contains outliers or shot noise, their estimation will degrade, as will be shown in our simulation results. To overcome the shortcomings of MCGF and VBGF, we take the advantages of VB and MCC by the fixed-point iteration method, and design the so called VBMCGF algorithm, which is summarized as follows:

Step 1: Predict:

$\hat{\mathbf{x}}_i^-$ and \mathbf{P}_i^- are obtained through (5) and (6), and

$$v_k^- = \rho(v_{k-1} - n - 1) + n + 1, \quad (38)$$

$$\mathbf{V}_k^- = \mathbf{B} \mathbf{V}_{k-1} \mathbf{B}^T, \quad (39)$$

where $0 < \rho \leq 1$, $0 < |\mathbf{B}| \leq 1$, and a reasonable choice is $\mathbf{B} = \sqrt{\rho} \mathbf{I}$.

Step 2: Update:

First, set $\hat{\mathbf{x}}_i^{(1)} = \hat{\mathbf{x}}_i^-$, $\mathbf{P}_i^{(1)} = \mathbf{P}_i^-$, $v_i = 1 + v_i^-$, and $\mathbf{V}_i^{(1)} = \mathbf{V}_i^-$. Calculate $\hat{\mathbf{z}}_i$ and \mathbf{P}_{xz} by (19) and (21). For $j = 1, \dots, N$, iterate the following equations:

$$\mathbf{R}_i^{(j)} = \mathbf{S}_{r,i}^{(j)} (\mathbf{S}_{r,i}^{(j)})^T = (v_i - n - 1)^{-1} \mathbf{V}_i^{(j)}, \quad (40)$$

$$\mathbf{e}_{r,i}^{(j)} = (\mathbf{S}_{r,i}^{(j)})^{-1} (\mathbf{z}_i - \hat{\mathbf{z}}_i), \quad (41)$$

$$\tilde{\mathbf{C}}_{r,i}^{(j)} = \text{diag} \left(G_\sigma(e_{r,i}^{(j)}(1)), \dots, G_\sigma(e_{r,i}^{(j)}(m)) \right), \quad (42)$$

$$\tilde{\mathbf{R}}_i^{(j)} = \mathbf{S}_{r,i}^{(j)} (\tilde{\mathbf{C}}_{r,i}^{(j)})^{-1} (\mathbf{S}_{r,i}^{(j)})^T, \quad (43)$$

$$\mathbf{T}_i^{(j+1)} = \int (\mathbf{h}_i(\mathbf{x}_i) - \hat{\mathbf{z}}_i)(\mathbf{h}_i(\mathbf{x}_i) - \hat{\mathbf{z}}_i)^T \mathcal{N}(\mathbf{x}_i | \hat{\mathbf{x}}_i^-, \mathbf{P}_i^-) d\mathbf{x}_i + \tilde{\mathbf{R}}_i^{(j)}, \quad (44)$$

$$\mathbf{K}_i^{(j+1)} = \mathbf{P}_{xz} (\mathbf{T}_i^{(j+1)})^{-1}, \quad (45)$$

$$\hat{\mathbf{x}}_i^{(j+1)} = \hat{\mathbf{x}}_i^- + \mathbf{K}_i^{(j+1)} (\mathbf{z}_i - \hat{\mathbf{z}}_i), \quad (46)$$

$$\mathbf{P}_i^{(j+1)} = \mathbf{P}_i^- - \mathbf{K}_i^{(j+1)} \mathbf{T}_i^{(j+1)} (\mathbf{K}_i^{(j+1)})^T, \quad (47)$$

$$\mathbf{V}_i^{(j+1)} = \mathbf{V}_i^- + \int (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i)) (\mathbf{z}_i - \mathbf{h}_i(\mathbf{x}_i))^T \mathcal{N}(\mathbf{x}_i | \hat{\mathbf{x}}_i^{(j+1)}, \mathbf{P}_i^{(j+1)}) d\mathbf{x}_i, \quad (48)$$

End For. In addition, set $\mathbf{V}_i = \mathbf{V}_i^{(N+1)}$, $\hat{\mathbf{x}}_i = \hat{\mathbf{x}}_i^{(N+1)}$, and $\mathbf{P}_i = \mathbf{P}_i^{(N+1)}$.

The main difference between the proposed VBMCGF and existing GFs lies in the modified estimation error covariance $\tilde{\mathbf{R}}_i^{(j)}$, where VB iterations are used to estimate its value and MCC is used to modify it in the presence of non-Gaussian noises. The kernel bandwidth σ plays an important role in reducing the effect of non-Gaussian noise or outliers. A smaller σ will make the filter more sensitive to outliers, but it may affect the convergence performance. In addition, a too large σ may cause the VBMCGF to perform more like VBGF (It can be proved that, if $\sigma \rightarrow \infty$, the proposed VBMCGF will reduce to VBGF). One possible way to select it is by the trial and error method [24–26]. Another important issue is the number of fixed-point iterations. In fact, only a few iterations (e.g., 2 or 3) are enough [31,32].

As the general framework, our filter can be easily implemented according to the real requirements. For linear systems that are described by $\mathbf{x}_i = \mathbf{F}_{i-1} \mathbf{x}_{i-1} + \mathbf{w}_{i-1}$ and $\mathbf{z}_i = \mathbf{H}_i \mathbf{x}_i + \mathbf{v}_i$, the predation update in the VBMCKF is the same as KF:

$$\hat{\mathbf{x}}_i^- = \mathbf{F}_{i-1} \hat{\mathbf{x}}_{i-1}, \quad (49)$$

$$\mathbf{P}_i^- = \mathbf{F}_{i-1} \mathbf{P}_{i-1} \mathbf{F}_{i-1}^T + \mathbf{Q}_{i-1}. \quad (50)$$

In addition, the $\hat{\mathbf{z}}_i$, \mathbf{P}_{xz} , $\mathbf{T}_i^{(j+1)}$, and $\mathbf{V}_i^{(j+1)}$ that appeared in VBMCKF will reduce to the following equations:

$$\hat{\mathbf{z}}_i = \mathbf{H}_i \hat{\mathbf{x}}_i^-, \quad (51)$$

$$\mathbf{P}_{xz} = \mathbf{P}_i^- \mathbf{H}_i^T, \quad (52)$$

$$\mathbf{T}_i^{(j+1)} = \mathbf{H}_i \mathbf{P}_i^- \mathbf{H}_i^T + \tilde{\mathbf{R}}_i^{(j)}, \quad (53)$$

$$\mathbf{V}_i^{(j+1)} = \mathbf{V}_i^- + \mathbf{H}_i \mathbf{P}_i^{(j+1)} \mathbf{H}_i^T + (\mathbf{z}_i - \mathbf{H}_i \hat{\mathbf{x}}_i^{(j+1)}) (\mathbf{z}_i - \mathbf{H}_i \hat{\mathbf{x}}_i^{(j+1)})^T, \quad (54)$$

while other steps are the same as the general framework.

When it comes to the nonlinear systems, the Gaussian integrals contained in $\hat{\mathbf{x}}_i^-$, \mathbf{P}_i^- , $\hat{\mathbf{z}}_i$, \mathbf{P}_{xz} , $\mathbf{T}_i^{(j+1)}$, and $\mathbf{V}_i^{(j+1)}$ can be calculated according to Taylor series, unscented transform, or cubature rules, and the corresponding filters are called VBMCEKF, VBMCKUF, and VBMCKKF, respectively.

4. Experimental Results

4.1. Simulation Results of the Target Tracking Model

To illustrate the performance of the proposed algorithm, we first give the simulation results using a typical target tracking model, where cubature rules are used to calculate the integrals. We compare the estimation accuracy of seven filters: CKF [7], MCCKF-1 [26], MCCKF-2 [25], VBCKF [9], HCKF [18], VBHCKF (which adopts Huber's function) and the proposed VBMCKF under various kinds of measurement noise. The target tracking example is modeled as [2]:

$$\mathbf{x}_i = \begin{bmatrix} 1 & T & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{i-1} + \begin{bmatrix} \frac{T^2}{2} & 0 \\ T & 0 \\ 0 & \frac{T^2}{2} \\ 0 & T \end{bmatrix} \mathbf{w}_{i-1}, \quad (55)$$

$$\mathbf{z}_i = \begin{bmatrix} \sqrt{(\zeta_x - \zeta_0)^2 + (\eta_y - \eta_0)^2} \\ \arctan \frac{\eta_x - \eta_0}{\zeta_y - \zeta_0} \end{bmatrix} + \mathbf{v}_i, \quad (56)$$

where $\mathbf{x}_i = [\zeta_{x,i} \ \dot{\zeta}_{x,i} \ \eta_{y,i} \ \dot{\eta}_{y,i}]^T$ is the system state with $(\zeta_{x,i}, \eta_{y,i})$ and $(\dot{\zeta}_{x,i}, \dot{\eta}_{y,i})$ being the position and velocity in x and y directions. The position of radar is set as $(\zeta_0, \eta_0) = (-100 \text{ m}, -100 \text{ m})$. We set the sampling period $T = 0.1 \text{ s}$, $\mathbf{Q}_i = \text{diag}(0.04 \text{ m}^2\text{s}^{-3}, 0.04 \text{ s}^{-3})$. The initial state is $\hat{\mathbf{x}}_0 = \bar{\mathbf{x}}_0 = [-40 \text{ m} \ 3 \text{ ms}^{-1} \ 10 \text{ m} \ 1 \text{ ms}^{-1}]^T$ and the covariance is $\mathbf{P}_0 = \text{diag}(4 \text{ m}^2, 0.01 \text{ m}^2\text{s}^{-2}, 4 \text{ m}^2, 0.01 \text{ m}^2\text{s}^{-2})$. The root mean square error (RMSE) and average RMSE (ARMSE) in position or velocity are used to describe the estimation accuracy—for example, the RMSE and ARMSE in position are defined as [2]:

$$\text{RMSE}_{pos}(i) = \sqrt{\frac{1}{M} \sum_{c=1}^M ((\zeta_{x,i}^c - \hat{\zeta}_{x,i}^c)^2 + (\eta_{y,i}^c - \hat{\eta}_{y,i}^c)^2)}, \quad (57)$$

$$\text{ARMSE}_{pos} = \frac{1}{L} \sum_{i=1}^L \text{RMSE}_{pos}(i), \quad (58)$$

where $(\zeta_{x,i}^c, \eta_{y,i}^c)$ and $(\hat{\zeta}_{x,i}^c, \hat{\eta}_{y,i}^c)$ are the true and estimated position in the c th Monte Carlo experiment, respectively. The RMSE and ARMSE of velocity are similar.

We here consider the following five kinds of measurement noises:

Case A: Gaussian distribution

$$\mathbf{v}_i \sim \mathcal{N}(\mathbf{0}, \text{diag}[(0.2 \text{ m})^2, (0.015 \text{ rad})^2]), \quad (59)$$

Case B: Time varying measurement noise covariance

$$\mathbf{v}_i \sim \mathcal{N}(\mathbf{0}, \alpha_i^2 \text{diag}[(0.2 \text{ m})^2, (0.015 \text{ rad})^2]), \quad (60)$$

Case C: Gaussian mixture noise with time varying measurement noise covariance

$$\mathbf{v}_i \sim 0.8\mathcal{N}(\mathbf{0}, \alpha_i^2 \text{diag}[(0.2 \text{ m})^2, (0.015 \text{ rad})^2]) + 0.2\mathcal{N}(\mathbf{0}, \text{diag}[(5 \text{ m})^2, (0.75 \text{ rad})^2]), \quad (61)$$

Case D: Time varying measurement noise covariance and shot noise

$$\mathbf{v}_i \sim [\beta_i, \gamma_i]^T + \mathcal{N}(\mathbf{0}, \alpha_i^2 \text{diag}[(0.2 \text{ m})^2, (0.015 \text{ rad})^2]), \quad (62)$$

Case E: Gaussian mixture noise with time varying measurement noise covariance and shot noise

$$\mathbf{v}_i \sim [\beta_i, \gamma_i]^T + 0.8\mathcal{N}(\mathbf{0}, \alpha_i^2 \text{diag}[(0.2 \text{ m})^2, (0.015 \text{ rad})^2]) + 0.2\mathcal{N}(\mathbf{0}, \text{diag}[(5 \text{ m})^2, (0.75 \text{ rad})^2]), \quad (63)$$

where the parameters α_i , β_i , and γ_i are given in Figure 1.

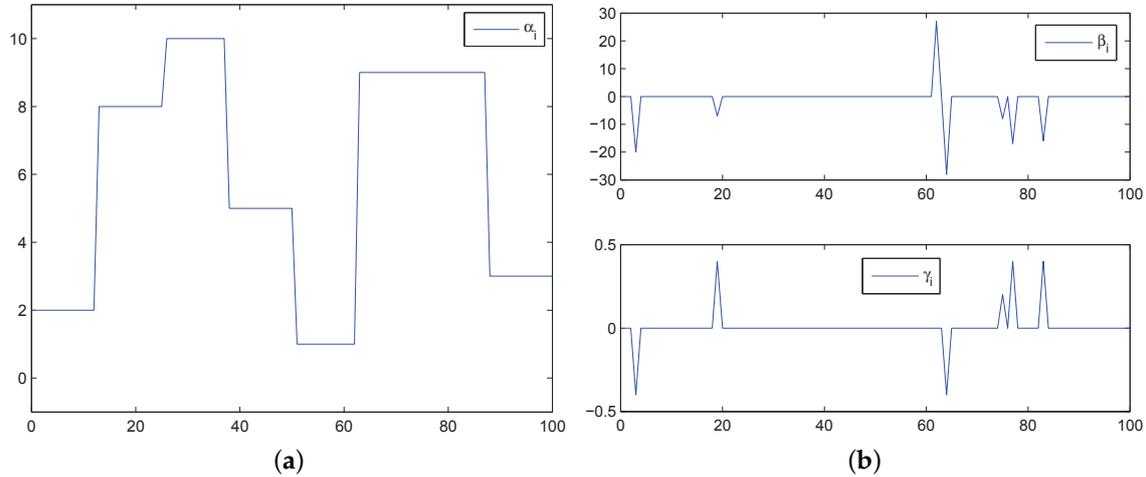


Figure 1. The time varying parameters. (a) α_i ; (b) β_i and γ_i .

In these simulations, we use $\sigma = 8$ for MCC based algorithms, the commonly used threshold $h = 1.345$ for Huber's function, and we set $\rho = 0.8$ and $N = 3$ for VB approximations. We perform $L = 100$ steps and run $M = 100$ Monte Carlo experiments for each case. The simulation results are plotted in Figures 2–6.

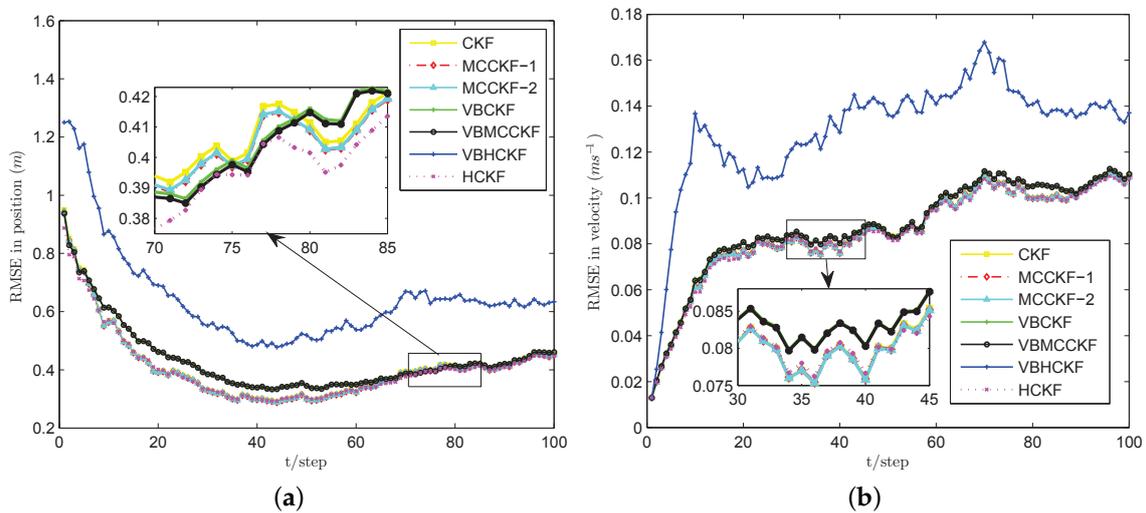


Figure 2. RMSE performances of different filters under Case A. (a) position; (b) velocity.

Under the Gaussian measurement noise with known noise covariance, as given in Figure 2, both MCCKF and HCKF have nearly similar estimation accuracy to CKF, since they will reduce to CKF if choosing proper free parameters (e.g., the σ and h are infinity). VBCKF and VBMCCKF work slightly worse as compared with CKF because they only use their online estimated measurement noise covariance instead of the real one. In particular, the VBHCKF has the worst performance since the commonly used parameter $h = 1.345$ for Huber's function doesn't fit the Gaussian noise situation when using the inaccurate online estimated measurement noise covariance. The proposed VBMCCKF

works well with the same kernel bandwidth under both Gaussian and non-Gaussian noise situations, as will be shown in the following cases.

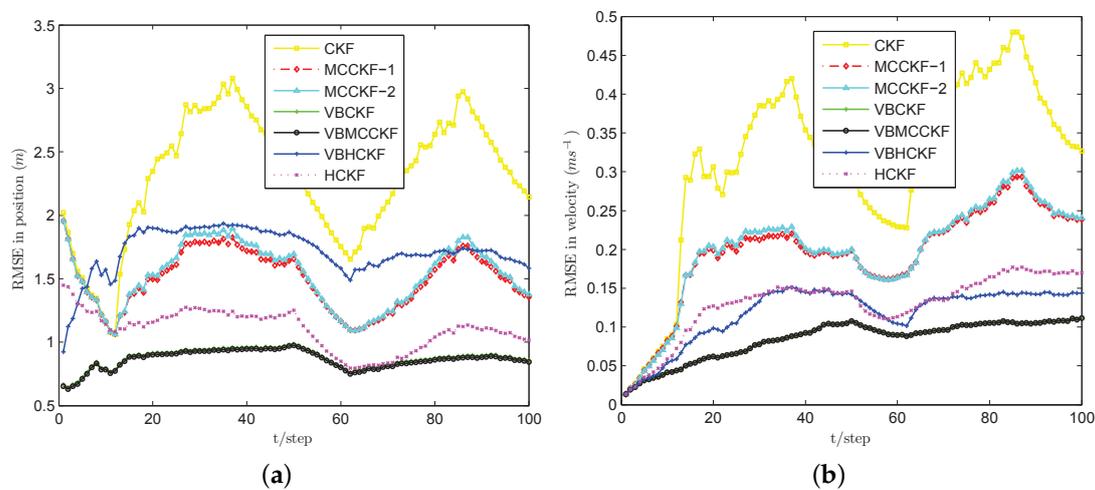


Figure 3. RMSE performances of different filters under Case B. (a) position; (b) velocity.

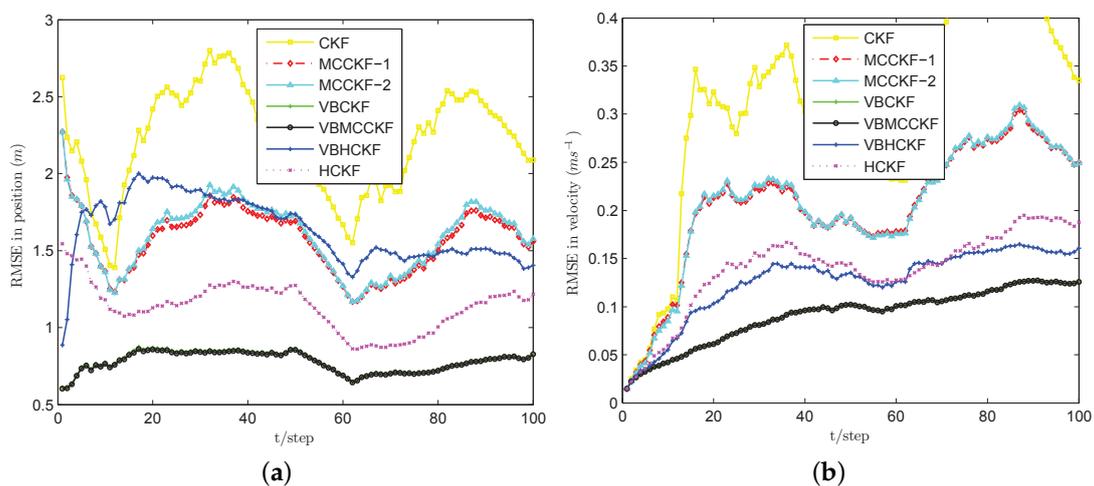


Figure 4. RMSE performances of different filters under Case C. (a) Position; (b) Velocity.

Figures 3–6 show the estimation performances of different algorithms under Cases B–E. It can be seen obviously that CKF has the worst estimation accuracy since it requires the measurement noise satisfying Gaussian distribution with known covariance, which is violated in these situations. MCCKF-1 and MCCKF-2 have similar estimation performance but are slightly worse than HCKF when using this kernel bandwidth. As demonstrated in [23–27], MCCKF is able to obtain better estimation accuracy than HCKF with a suitable σ . The estimation results of MCCKF and HCKF do not change too much when Gaussian mixture noise or shot noise are added, since they are robust filters. The VBHCKF has better estimation results in velocity but worse accuracy than HCKF. The VBCKF has much better estimation in Cases B and C, as it is able to online estimate the time varying measurement noise covariance. However, its performance will degrade once shot noise is injected in Cases D and E. Among these algorithms, our VBMCKKF has the best estimation accuracy as compared with other algorithms under Cases B–E. It shows the adaptivity to unknown changing measurement noise covariance and robustness to Gaussian mixture noise and shot noise. Its estimation results are also much better than VBHCKF since the MCC has the potential to capture high order information than Huber’s function. The ARMSEs of these filters under different noises are also given in Table 1 to clearly show the differences.

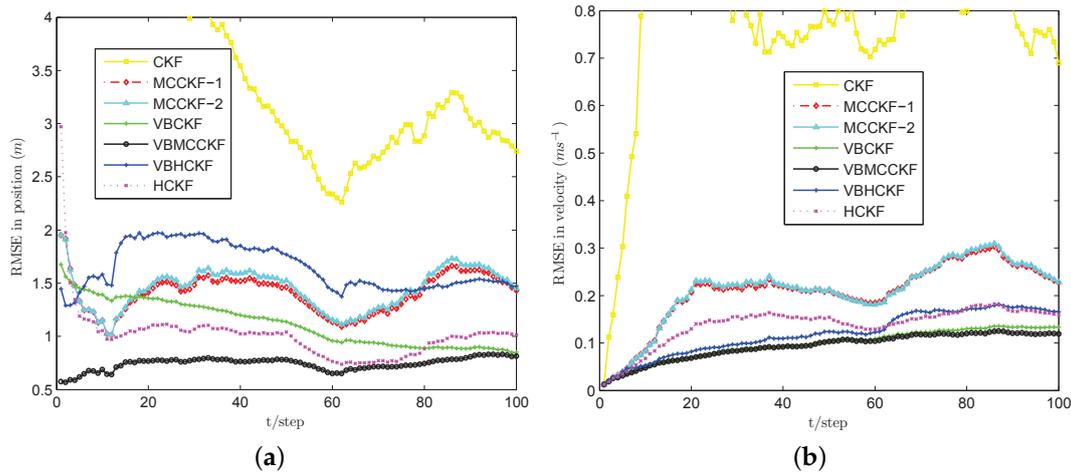


Figure 5. RMSE performances of different filters under Case D. (a) position; (b) velocity.

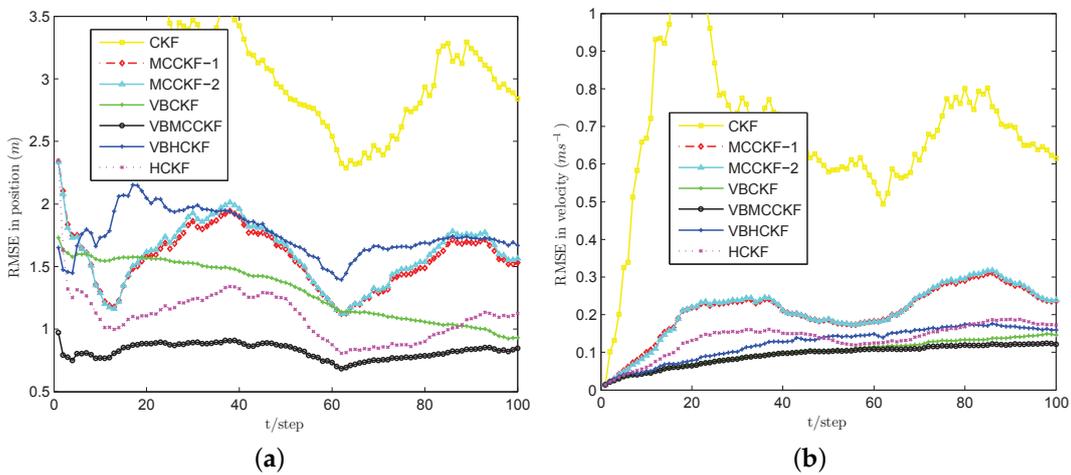


Figure 6. RMSE performances of different filters under Case E. (a) position; (b) velocity.

Table 1. ARMSEs of different filters under five cases.

Algorithms	Case A		Case B		Case C		Case D		Case E	
	Pos.	Vel.								
CKF	0.4097	0.0855	2.2930	0.3121	2.2050	0.3053	3.7720	0.7809	3.2960	0.6923
MCCKF-1	0.4077	0.0854	1.4960	0.1959	1.5680	0.2052	1.4090	0.2077	1.5680	0.2100
MCCKF-2	0.4079	0.0853	1.5280	0.1989	1.5990	0.2058	1.4500	0.2098	1.6010	0.2114
HCKF	0.4045	0.0847	1.0990	0.1304	1.1360	0.1390	1.0100	0.1384	1.1030	0.1365
VBHCKF	0.6500	0.1296	1.7180	0.1185	1.6360	0.1260	1.6460	0.1207	1.7590	0.1251
VBCKF	0.4362	0.0880	0.8761	0.0845	0.7828	0.0914	1.1240	0.0977	1.3100	0.1000
VBMCKKF	0.4350	0.0878	0.8633	0.0843	0.7752	0.0909	0.7424	0.0933	0.8236	0.0931

4.2. Field Results of Integrated Navigation

To further illustrate the effectiveness of the proposed algorithm, we compare our algorithm and existing related methods using the real data collected by a self-made fiber optical gyroscope inertial navigation system (INS) together with a doppler velocity logger (DVL). The integrated navigation results of photonics inertial navigation system (PHINS) and GPS are used as the reference system. We adopt the loosely coupled method to fusion the information of INS and DVL. The state vector is chosen as $x = [\delta L \ \delta \lambda \ \delta V_E \ \delta V_N \ \varphi_x \ \varphi_y \ \varphi_z \ \nabla_x \ \nabla_y \ \nabla_z \ \varepsilon_x \ \varepsilon_y \ \varepsilon_z]^T$, where δL and $\delta \lambda$ are the latitude and longitude error, $\{\delta V_j, \varphi_j, \nabla_j, \varepsilon_j\}$ are the velocity error, attitude error, accelerometer bias and gyroscope

constant drift, respectively. j denotes the subscribe $\{e, n, x, y, z\}$, where e and n present the east and north directions in the local-level frame, and $x, y,$ and z are the directions of three axes in the body frame. Then, the continuous system model is given as follows:

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{w}(t), \quad (64)$$

where t is the continuous system time, and $\mathbf{w}(t) = [\mathbf{0}_{1 \times 2} \ w_{ax} \ w_{ay} \ w_{gx} \ w_{gy} \ w_{gz} \ \mathbf{0}_{1 \times 5}]^T$ is the process noise, which contains the Gaussian noise of both accelerometers and gyroscopes. The detailed elements of matrix $\mathbf{A}(t)$ and $\mathbf{B}(t)$ can refer to [34]. The measurement equation is

$$\mathbf{z}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{v}(t), \quad (65)$$

where $\mathbf{z}(t) = [(V_e^{INS} - V_e^{DVL}) \ (V_n^{INS} - V_n^{DVL})]^T$, $\mathbf{H} = [\mathbf{0}_{2 \times 2} \ \mathbf{I}_{2 \times 2} \ \mathbf{0}_{2 \times 8}]$, and $\mathbf{v}(t) = [V_e^{DVL} \ V_n^{DVL}]^T$. Then, the discretization process is performed before running filtering algorithms.

We compare the estimation results of KF, VBKF, HKF, MCKF, and VBMCKF, where we choose $h = 1.345$, $\sigma = 4$, $\rho = 0.96$, and $N = 3$. We set $\hat{\mathbf{x}}_0 = [\mathbf{0}_{1 \times 12}]^T$, and the covariance is $\mathbf{P}_0 = \text{diag}(\frac{500}{R_e} \text{ rad/s}, \frac{500}{R_e} \text{ rad/s}, 0.5^\circ, 0.5^\circ, 3^\circ, 1 \times 10^{-4} \text{ g}, 1 \times 10^{-4} \text{ g}, 0.01^\circ/\text{h}, 0.01^\circ/\text{h}, 0.01^\circ/\text{h})$, where R_e is the radius of the earth, and g is the gravitational acceleration. The process noise covariance and measurement noise covariance are set as $\mathbf{Q}_i = \text{diag}(0, 0, 1 \times 10^{-3} \text{ g}, 1 \times 10^{-3} \text{ g}, 0.025^\circ/\text{h}, 0.025^\circ/\text{h}, 0.025^\circ/\text{h}, 0, 0, 0, 0, 0, 0)$ and $\mathbf{R}_i = \text{diag}(0.1 \text{ m/s}, 0.1 \text{ m/s})$ according to the parameters of INS and DVL.

In this experiment, the system is first tested in anchorage for about 50 min, then the ship starts to move. The real velocities of the ship are shown in Figure 7 provided by the commercial INS/GPS integrated navigation system. The collected data is processed using MATLAB (R2014a by MathWorks, Inc., Natick, MA, USA) on a computer with 2.50 GHz Intel Core i5-7300HQ CPU and 8 GB memory. The total computational time of KF, VBKF, HKF, MCKF, and VBMCKF are 0.1900 s, 0.4800 s, 0.2640 s, 0.2310 s, and 0.6340 s, respectively. The position and velocity errors of different filters are given in Figures 8 and 9. The differences between attitude and heading errors are quite similar so we omit them.

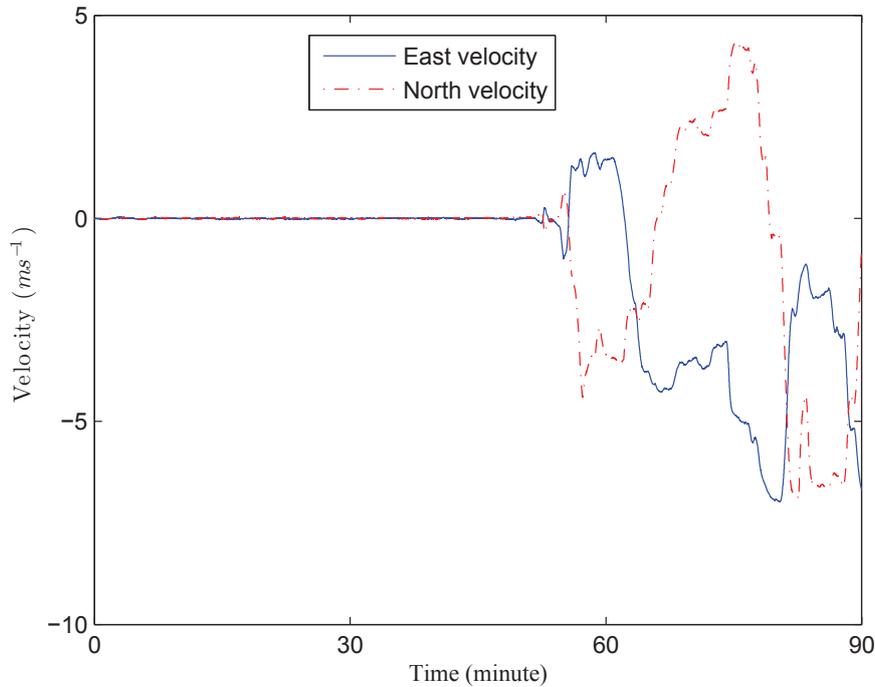


Figure 7. Real velocities of this field experiment.

It can be seen from Figures 8 and 9 that when the motion state changes sharply, the proposed VBMCKF algorithm has the smallest estimation errors with slightly increased computational time as compared with other estimation methods.

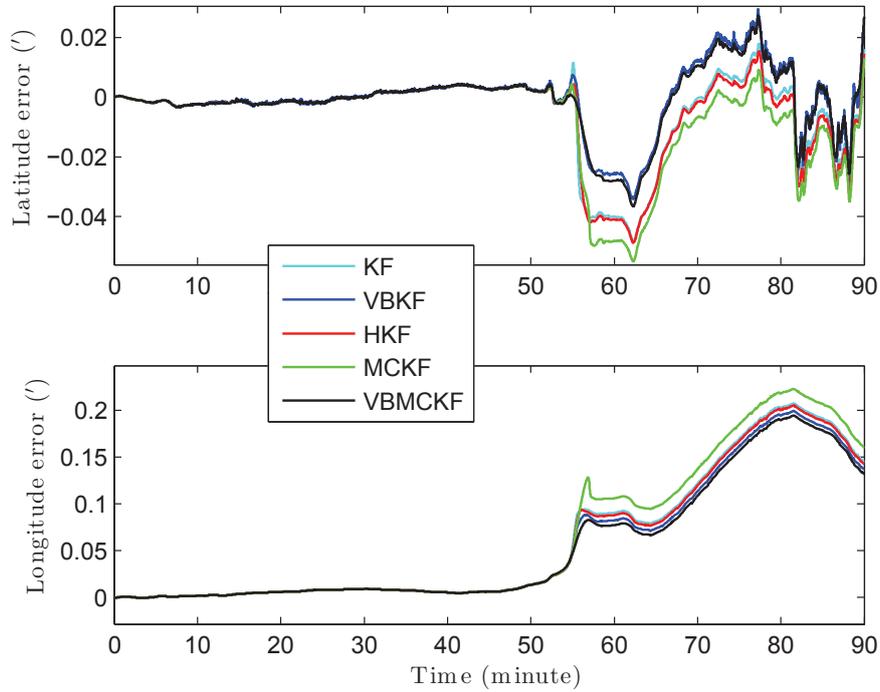


Figure 8. Position errors of different filters.

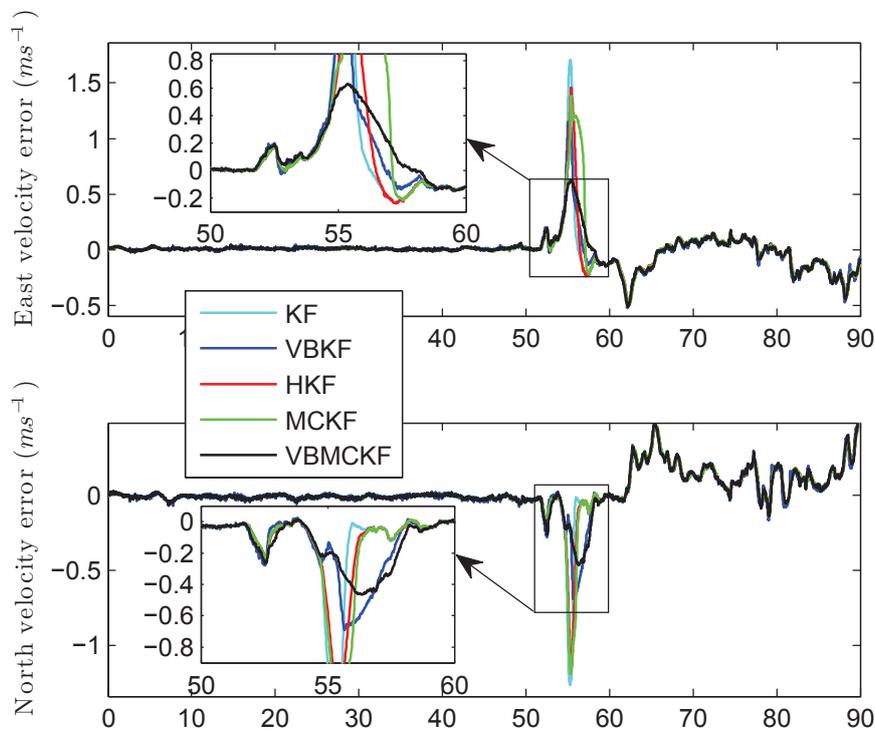


Figure 9. Velocity errors of different filters.

5. Conclusions

In this paper, a novel adaptive MCGF based on VB approximation is proposed. The MCC is used to reduce the effect of non-Gaussian measurement noise and outliers, while we use VB to estimate the unknown measurement noise covariance. Experimental results based on simulation examples and real data show that the proposed algorithm has better estimation accuracy than related robust and adaptive filters.

Author Contributions: W.G. and G.Z. conceived and designed the experiments; W.G. and Z.Y. performed the experiments and analyzed the data; M.B. and Z.Y. contributed experiment tools; W.G. and G.Z. wrote the paper.

Funding: This research was funded by the National Natural Science Foundation of China under grant number 61773133 and 61633008, the Natural Science Foundation of Heilongjiang Province under grant number F2016008, the Fundamental Research Funds for the Central Universities under grant number HEUCFM180401, the China Scholarship Council Foundation, and the Ph.D. Student Research and Innovation Foundation of the Fundamental Research Funds for the Central Universities under grant number HEUGIP201807.

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

References

- Auger, F.; Hilaret, M.; Guerrero, J.M.; Monmasson, E.; Orłowska-Kowalska, T.; Katsura, S. Industrial Applications of the Kalman Filter: A Review. *IEEE Trans. Ind. Electron.* **2013**, *60*, 5458–5471.
- Wang, G.; Li, N.; Zhang, Y. An Event based Multi-sensor Fusion Algorithm with Deadzone Like Measurements. *Inf. Fusion* **2018**, *42*, 111–118.
- Wang, G.; Li, N.; Zhang, Y. Diffusion Distributed Kalman Filter over Sensor Networks without Exchanging Raw Measurements. *Signal Process.* **2017**, *132*, 1–7.
- Huang, Y.; Zhang, Y.; Xu, B.; Wu, Z.; Chambers, J. A New Outlier-Robust Student's t Based Gaussian Approximate Filter for Cooperative Localization. *IEEE ASME Trans. Mechatron.* **2017**, *22*, 2380–2386.
- Reif, K.; Günther, S.; Yaz, E.; Unbehauen, R. Stochastic Stability of the Discrete-time Extended Kalman Filter. *IEEE Trans. Autom. Control* **1999**, *44*, 714–728.
- Julier, S.J.; Uhlmann, J.K. Unscented Filtering and Nonlinear Estimation. *Proc. IEEE* **2004**, *92*, 401–422.
- Arasaratnam, I.; Haykin, S. Cubature Kalman Filters. *IEEE Trans. Autom. Control* **2009**, *54*, 1254–1269.
- N Rgaard, M.; Poulsen, N.K.; Ravn, O. New Developments in State Estimation for Nonlinear Systems. *Automatica* **2000**, *36*, 1627–1638.
- Särkkä, S.; Hartikainen, J. Non-linear Noise Adaptive Kalman Filtering via Variational Bayes. In Proceedings of the 2013 IEEE International Workshop on Machine Learning for Signal Processing, Southampton, UK, 14 November 2013; pp. 1–6.
- Huang, Y.; Zhang, Y.; Wang, X.; Zhao, L. Gaussian Filter for Nonlinear Systems with Correlated Noises at the Same Epoch. *Automatica* **2015**, *60*, 122–126.
- Wang, G.; Li, N.; Zhang, Y. Hybrid Consensus Sigma Point Approximation Nonlinear Filter Using Statistical Linearization. *Trans. Inst. Meas. Control* **2018**, *40*, 2517–2525.
- Shmaliy, Y.S.; Zhao, S.; Ahn, C.K. Unbiased Finite Impulse Response Filtering: An Iterative Alternative to Kalman Filtering Ignoring Noise and Initial Conditions. *IEEE Control Syst. Mag.* **2017**, *37*, 70–89.
- Zhao, S.; Shmaliy, Y.S.; Shi, P.; Ahn, C.K. Fusion Kalman/UFIR Filter for State Estimation with Uncertain Parameters and Noise Statistics. *IEEE Trans. Ind. Electron.* **2017**, *64*, 3075–3083.
- Kotecha, J.H.; Djuric, P.M. Gaussian Particle Filtering. *IEEE Trans. Signal Process.* **2003**, *51*, 2592–2601.
- Pak, J.M.; Ahn, C.K.; Shmaliy, Y.S.; Lim, M.T. Improving Reliability of Particle Filter-Based Localization in Wireless Sensor Networks via Hybrid Particle/FIR Filtering. *IEEE Trans. Ind. Inform.* **2015**, *11*, 1089–1098.
- Alspach, D.; Sorenson, H. Nonlinear Bayesian Estimation Using Gaussian Sum Approximations. *IEEE Trans. Autom. Control* **1972**, *17*, 439–448.
- Li, X.R.; Jilkov, V.P. Survey of Maneuvering Target Tracking. Part V. Multiple-model Methods. *IEEE Trans. Aerosp. Electron. Syst.* **2005**, *41*, 1255–1321.
- Chang, L.; Hu, B.; Chang, G.; Li, A. Multiple Outliers Suppression Derivative-Free Filter Based on Unscented Transformation. *J. Guid. Control Dyn.* **2012**, *35*, 1902–1907.

19. Karlgaard, C.D. Nonlinear Regression Huber-Kalman Filtering and Fixed-Interval Smoothing. *J. Guid. Control Dyn.* **2014**, *38*, 322–330.
20. Wu, H.; Chen, S.; Yang, B.; Chen, K. Robust Derivative-Free Cubature Kalman Filter for Bearings-Only Tracking. *J. Guid. Control Dyn.* **2016**, *39*, 1866–1871.
21. Hur, H.; Ahn, H.S. Discrete-Time H_∞ Filtering for Mobile Robot Localization Using Wireless Sensor Network. *IEEE Sens. J.* **2013**, *13*, 245–252.
22. Graham, M.C.; How, J.P.; Gustafson, D.E. Robust State Estimation with Sparse Outliers. *J. Guid. Control Dyn.* **2015**, *38*, 1229–1240.
23. Wang, Y.; Zheng, W.; Sun, S.; Li, L. Robust Information Filter Based on Maximum Correntropy Criterion. *J. Guid. Control Dyn.* **2016**, *39*, 1126–1131.
24. Chen, B.; Liu, X.; Zhao, H.; Principe, J.C. Maximum Correntropy Kalman Filter. *Automatica* **2017**, *76*, 70–77.
25. Liu, X.; Qu, H.; Zhao, J.; Yue, P.; Wang, M. Maximum Correntropy Unscented Kalman Filter for Spacecraft Relative State Estimation. *Sensors* **2016**, *16*, 1530.
26. Liu, X.; Chen, B.D.; Xu, B.; Wu, Z.Z.; Honeine, P. Maximum Correntropy Unscented Filter. *Int. J. Syst. Sci.* **2017**, *48*, 1607–1615.
27. Liu, X.; Qu, H.; Zhao, J.; Chen, B. Extended Kalman Filter under Maximum Correntropy Criterion. In Proceedings of the 2016 International Joint Conference on Neural Networks, Vancouver, BC, Canada, 24–29 July 2016; pp. 1733–1737.
28. Wang, G.; Li, N.; Zhang, Y. Maximum Correntropy Unscented Kalman and Information Filters for Non-Gaussian Measurement Noise. *J. Frankl. Inst.* **2017**, *354*, 8659–8677.
29. Narasimhappa, M.; Rangababu, P.; Sabat, S.L.; Nayak, J. A modified Sage–Husa Adaptive Kalman Filter for Denoising Fiber Optic Gyroscope Signal. In Proceedings of the 2012 Annual IEEE India Conference, Kochi, India, 7–9 December 2012; pp. 1266–1271.
30. Wang, Y.; Sun, S.; Li, L. Adaptively Robust Unscented Kalman Filter for Tracking a Maneuvering Vehicle. *J. Guid. Control Dyn.* **2014**, *37*, 1696–1701.
31. 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.
32. Sarkka, S.; Nummenmaa, A. Recursive Noise Adaptive Kalman Filtering by Variational Bayesian Approximations. *IEEE Trans. Autom. Control* **2009**, *54*, 596–600.
33. Mbalawata, I.S.; Särkkä, S.; Vihola, M.; Haario, H. Adaptive Metropolis Algorithm using Variational Bayesian Adaptive Kalman Filter. *Comput. Stat. Data Anal.* **2015**, *83*, 101–115.
34. Gao, W.; Li, J.; Zhou, G.; Li, Q. Adaptive Kalman Filtering with Recursive Noise Estimator for Integrated SINS/DVL Systems. *J. Navig.* **2015**, *68*, 142–161.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).