

## Research Article

# Support Vector Regression-Based Adaptive Divided Difference Filter for Nonlinear State Estimation Problems

**Hongjian Wang, Jinlong Xu, Aihua Zhang, Cun Li, and Hongfei Yao**

*College of Automation, Harbin Engineering University, Harbin 150001, China*

Correspondence should be addressed to Jinlong Xu; [xujinlong\\_1983@sina.cn](mailto:xujinlong_1983@sina.cn)

Received 2 March 2014; Accepted 4 May 2014; Published 25 May 2014

Academic Editor: Weichao Sun

Copyright © 2014 Hongjian Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a support vector regression-based adaptive divided difference filter (SVRADDf) algorithm for improving the low state estimation accuracy of nonlinear systems, which are typically affected by large initial estimation errors and imprecise prior knowledge of process and measurement noises. The derivative-free SVRADDf algorithm is significantly simpler to compute than other methods and is implemented using only functional evaluations. The SVRADDf algorithm involves the use of the theoretical and actual covariance of the innovation sequence. Support vector regression (SVR) is employed to generate the adaptive factor to tune the noise covariance at each sampling instant when the measurement update step executes, which improves the algorithm's robustness. The performance of the proposed algorithm is evaluated by estimating states for (i) an underwater nonmaneuvering target bearing-only tracking system and (ii) maneuvering target bearing-only tracking in an air-traffic control system. The simulation results show that the proposed SVRADDf algorithm exhibits better performance when compared with a traditional DDF algorithm.

## 1. Introduction

The problem of state estimation for nonlinear systems has been a subject of considerable research interest in recent years, but there is still no single solution that outperforms all other approaches. Most proposed estimators are nonlinear extensions of the dominated Kalman filter (see [1]), and each approach provides a suboptimal trade-off between properties such as numerical robustness, computational burden, and estimation accuracy. The extended Kalman filter (EKF), which linearizes both nonlinear terms of a current estimated state trajectory, is based on a first-order Taylor series and displays poor performance if the system is highly nonlinear. The limitations of EKFs are enumerated in [2]. Another improved algorithm is the iterated extended Kalman filter (IEKF), which linearizes the nonlinear model around an updated state rather than the predicted state (see [3]). Although IEKFs have been proven to perform better than EKFs in addition to globally guaranteeing convergence, the algorithm still requires a Jacobian matrix just like EKFs. However, no solution exists for the Jacobian matrix in nonlinear systems

for some situations, which limits the potential application of both EKFs and IEKFs.

In recent years, a new class of filter known as sigma-point Kalman filter (SPKF) has attracted a great deal of attention. In SPKFs, the algorithm propagates a cluster of points centered on the current state instead of linearizing the system dynamics to improve the approximations of the conditional mean and covariance. Unscented Kalman filters (UKF) and divided difference filters (DDF) are two kinds of SPKFs.

UKFs use a deterministic sampling technique to pick a minimal set of sample points around the mean to catch the higher order statistics of the system so as to better estimation accuracy and convergence characteristics (see [4]). In [5, 6], a UKF for a class of nonlinear discrete-time systems with correlated noises was designed to deal with the problem of nonlinear filtering failure found in conventional UKFs when system noise is correlated with measurement noise. The proposed UKF breaks the limitation of conventional UKFs that requires system noise and measurement noise to be uncorrelated Gauss white noises, thus extending the potential

application of conventional UKFs. In [7], a UKF filtering algorithm with colored measurement noise was proposed. The algorithm was first derived on the basis of augmented measurement information and minimum mean square error estimation, and a filtering recursive formula of UKF with colored noise then added by applying an unscented transformation to calculate the posterior mean and covariance of the nonlinear state within the optimal framework. The proposed UKF effectively dealt with the fact that traditional UKFs fail when measurement noise is colored. In [8], a UKF was applied to multiple target tracking, with the proposed UKF shown to have improved performance versus previous EKF approaches. In [9], the paper discussed an adaptive multiuser receiver for CDMA systems in which the scaled unscented filter (SUF) and the square root unscented filter (SURF) were used for joint estimation and tracking of the code delays and multipath coefficients of the received CDMA signals. The proposed channel estimators were more near-far resistant than in conventional EKFs and presented lower complexity than conventional particle filter- (PF-) based methods. Computer simulation results demonstrated the superior performance of the proposed channel estimators, and the proposed estimators were shown to exhibit lower complexity relative to the PF-based method. Although UKFs have undergone a significant amount of meaningful theory innovation and are now used in many fields, [10–12] show that UKF accuracy is lower than that of DDF, while also having a higher computational cost.

The divided difference filter (DDF) first proposed by Nøgaard et al. (see [13]) linearizes the nonlinear terms based on Stirling's interpolation polynomial approximations formula rather than Taylor's approximation of nonlinear terms in an EKF. Conceptually, the implementation principle resembles that of an EKF; however, the DDF is significantly simpler as it does not need to calculate the Jacobian matrix and no derivatives are required. The DDF that Nøgaard et al. developed works on general discrete-time nonlinear models in which the noises are not assumed to be additive. In [14], the paper further formulated a DDF in terms of the innovation vector approach, the additive process, and the measurement noise sources. In [15], the paper proposed a new filter named the maximum likelihood-based iterated divided difference filter (MLIDDF), which improved the low state estimation accuracy of nonlinear state estimation that results from large initial estimation errors and the nonlinearity of the measurement equations. Simulation results showed that the MLIDDF algorithm possessed better state estimation accuracy and a faster convergence rate. In [16], the authors proposed a novel adaptive version of the DDF that was applicable to nonlinear systems with a linear output equation. In order to make the filter robust to modeling errors, upper bounds on the state covariance matrix were derived. The parameters of the upper bound were then estimated using a combination of offline tuning and online optimization with a linear matrix inequality constraint, which ensured that the predicted output error covariance was larger than the observed output error covariance. Simulation results demonstrated the superior performance of the proposed filter as compared to the standard DDF. Reference [17] presented

an ensemble-based approach that handled nonlinearity based on a simplified divided difference approximation through Stirling's interpolation formula. The algorithm used Stirling's interpolation formula to evaluate the statistics of the background ensemble during the prediction step, employing the formula in an ensemble square root filter (EnSRF) at the filtering step to update the background for analysis. In this sense, the algorithm is a hybrid of Stirling's interpolation formula and the EnSRF method, while the computational cost of the algorithm is less than that of EnSRF.

Different studies have focused on the application of DDFs to nonlinear state estimation problems. In [18], time delay and channel gain estimation for multipath fading code division multiple access (CDMA) signals using a DDF were investigated, and the simulation results showed that the DDF was simpler to implement and more resilient to near-far interference in CDMA networks compared with an EKF. In [19, 20], the relative kinematic states of a reentry vehicle obtained from noisy seeker measurements using a DDF were examined. The results were compared to those obtained using an EKF and a UKF and showed that the DDF was more accurate than estimators based on a Taylor approximation like the EKF. Reference [21] investigated the possibility of using a DDF for estimating the internal variables of a synchronous generator, such as the rotor angle where the data acquired is from a phasor measurement. The effectiveness of the method was tested on a single machine infinite bus system, a nine-bus system, and a 68-bus New England-New York interconnected system. In [22], a DDF using orientation estimation was considered. The fourth element of the quaternion error vector was removed from the system states to alleviate estimated error covariance matrix divergence. The measurement system was a MARG sensor, which consisted of a triaxial rate gyro, a triaxial accelerometer, and a triaxial magnetometer. The nonlinear measurement model was obtained based on the principals of operation of the magnetometer and accelerometer and the properties of the quaternion vector space. The performance of three filters, DDF, EKF, and UKF, was compared with different sampling frequencies. The work showed that the tested DDF and the UKF were more robust than the EKF under the same initial angle-error conditions. The DDF also performed better than the UKF, although the computational load for the UKF was less. In [23], a DDF-based data fusion algorithm was presented, which utilized the complementary noise profile of rate gyros and gravimetric inclinometers to extend their limits and achieve more accurate attitude estimates. In [24], a DDF-based ballistic target tracking system for the reentry phase was proposed. The paper compared DDF, EKF, and UKF algorithms using a Monte Carlo simulation approach, with the simulation results showing that the DDF outperformed both the EKF and UKF in terms of estimation accuracy and filtering credibility. In [25], a DDF with quaternion-based dynamic process modeling was applied to global positioning system (GPS) navigation to increase navigation estimation accuracy at high-dynamic regions while preserving precision at low-dynamic regions. Some properties and performance metrics were assessed and compared to those using EKF and UKF approaches.

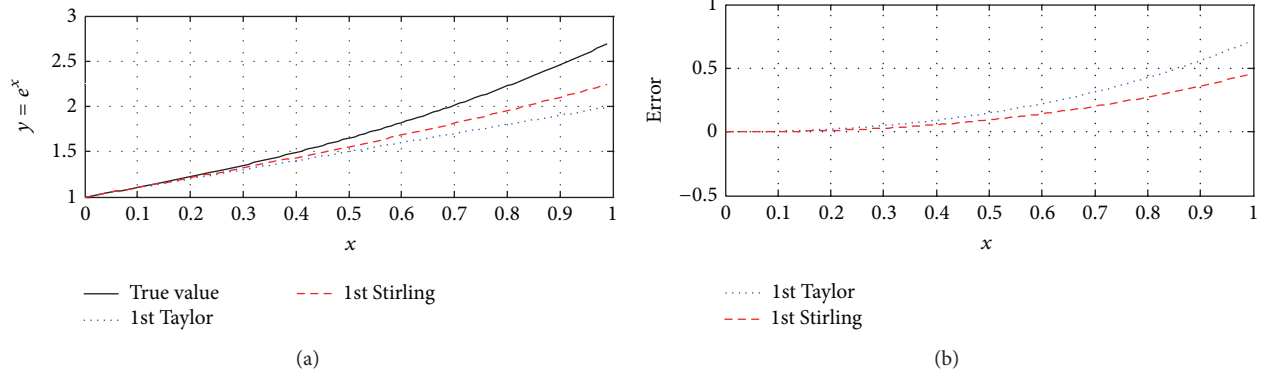


FIGURE 1: Comparison of first-order Stirling series with first-order Taylor series results.

Despite their recent popularity, DDF algorithms require that both the system model and the stochastic information must be accurate. However, this condition cannot be satisfied in many practical situations, which forces the filter to adapt itself to changing conditions. One of the problems with this requirement is that any change in the process introduces measurement noise covariance. In this work, we make use of the theoretical and actual covariance of the innovation sequence, employing SVR to generate the scale factor to tune the noise covariance at each sampling instant when the measurement update step is executed to adapt the filtering algorithm. This paper is organized as follows. Section 2 briefly introduces DDF theory and the proposed SVR-based adaptive strategy. Passive target tracking is then carried out to evaluate the performance of DDF and SVRADDF algorithms using a Monte Carlo simulation in Section 3. Finally, conclusions are provided in Section 4.

## 2. Development of the Support Vector Regression-Based Adaptive Divided Difference Filter

**2.1. Divided Difference Filter.** Consider the nonlinear function

$$\mathbf{y} = \mathbf{f}(\mathbf{x}), \quad (1)$$

where  $\mathbf{x} \in \mathbf{R}^{n_x}$  and  $\mathbf{y} \in \mathbf{R}^{n_y}$ . If the function is analytic, then the first-order Taylor series expanded about some point  $\mathbf{x} = \bar{\mathbf{x}}$  becomes

$$\begin{aligned} \mathbf{y} = & \left( \mathbf{f}(\bar{\mathbf{x}} + \Delta\mathbf{x}) = \mathbf{f}(\bar{\mathbf{x}}) + \mathbf{D}_{\Delta\mathbf{x}}\mathbf{f} = \mathbf{f}(\bar{\mathbf{x}}) \right. \\ & \left. + \left( \Delta x_1 \frac{\partial}{\partial x_1} + \Delta x_2 \frac{\partial}{\partial x_2} + \cdots + \Delta x_n \frac{\partial}{\partial x_n} \right) \mathbf{f}(\mathbf{x}) \right) \Big|_{\mathbf{x}=\bar{\mathbf{x}}} \end{aligned} \quad (2)$$

with (2) truncated after the first-order term. Note that (2) can achieve a better local approximation if more terms are included. However, such an expanded Taylor series requires derivatives and cannot be fulfilled in some situations. Stirling's interpolation formula is based on a finite number of

evaluations of the function and does not require derivatives, with the first-order approximation yielding

$$\begin{aligned} \mathbf{y} = \mathbf{f}(\bar{\mathbf{x}} + \Delta\mathbf{x}) &= \mathbf{f}(\bar{\mathbf{x}}) + \tilde{\mathbf{D}}_{\Delta\mathbf{x}}\mathbf{f}(\bar{\mathbf{x}}) \\ &= \mathbf{f}(\bar{\mathbf{x}}) + \frac{1}{h} \left( \sum_{i=1}^n \Delta x_i \mu_i \delta_i \right) \mathbf{f}(\bar{\mathbf{x}}), \end{aligned} \quad (3)$$

where  $h$  denotes a selected interval length and  $\delta$  and  $\mu$  are determined by

$$\begin{aligned} \delta_i \mathbf{f}(\bar{\mathbf{x}}) &= \mathbf{f}\left(\bar{\mathbf{x}} + \frac{h}{2}\mathbf{e}_i\right) - \mathbf{f}\left(\bar{\mathbf{x}} - \frac{h}{2}\mathbf{e}_i\right), \\ \mu_i \mathbf{f}(\bar{\mathbf{x}}) &= \frac{1}{2} \left( \mathbf{f}\left(\bar{\mathbf{x}} + \frac{h}{2}\mathbf{e}_i\right) + \mathbf{f}\left(\bar{\mathbf{x}} - \frac{h}{2}\mathbf{e}_i\right) \right), \end{aligned} \quad (4)$$

with  $\mathbf{e}_i$  being the  $i$ th unit vector.

Figure 1 compares the results found by using (2) and (3). The function example is  $f(x) = e^x$ , where  $h = 0.56$ . From the figure, we can see that Stirling's interpolation provides better accuracy than the Taylor series under the same order approximations.

We now assume that the variable  $\mathbf{x}$  has a Gaussian density with mean  $\bar{\mathbf{x}}$  and covariance  $\mathbf{P}_{\mathbf{x}}$ . We can introduce a transformation matrix  $\mathbf{S}_{\mathbf{x}}$  which we select as a square Cholesky factor of  $\mathbf{P}_{\mathbf{x}}$ , such that  $\mathbf{P}_{\mathbf{x}} = \mathbf{S}_{\mathbf{x}}\mathbf{S}_{\mathbf{x}}^T$ . To illustrate how others can be derived, we introduce the linear transformation of  $\mathbf{x}$ :

$$\mathbf{z} = \mathbf{S}_{\mathbf{x}}^{-1}\mathbf{x}. \quad (5)$$

This linear transformation results in a stochastic decoupling of  $\mathbf{x}$  as the elements of  $\mathbf{z}$  become mutually uncorrelated (see [13]). This changes (3) to

$$\mathbf{y} = \mathbf{f}(\mathbf{S}_{\mathbf{x}}\mathbf{z}) = \tilde{\mathbf{f}}(\mathbf{z}) = \tilde{\mathbf{f}}(\bar{\mathbf{z}}) + \tilde{\mathbf{D}}_{\Delta\mathbf{z}}\tilde{\mathbf{f}}, \quad (6)$$

where  $\tilde{\mathbf{D}}_{\Delta\mathbf{z}}\tilde{\mathbf{f}}$  is determined by

$$\tilde{\mathbf{D}}_{\Delta\mathbf{z}}\tilde{\mathbf{f}} = \frac{1}{h} \left( \sum_{i=1}^n \Delta z_i \mu_i \delta_i \right) \mathbf{f}(\bar{\mathbf{z}}). \quad (7)$$

The mean  $\bar{\mathbf{y}}$ , covariance  $\mathbf{P}_{yy}$ , and cross covariance  $\mathbf{P}_{xy}$  of  $\mathbf{y}$  are obtained from

$$\begin{aligned}\bar{\mathbf{y}} &= \mathbf{E}[\mathbf{y}] = \mathbf{E}[\tilde{\mathbf{f}}(\bar{\mathbf{z}}) + \tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}}] = \mathbf{E}[\tilde{\mathbf{f}}(\bar{\mathbf{z}})] = \mathbf{f}(\bar{\mathbf{x}}), \\ \mathbf{P}_{yy} &= \mathbf{E}[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})^T] \\ &= \mathbf{E}[(\tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}})(\tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}})^T] \\ &= \frac{1}{4h^2} \sum_{i=1}^n (\mathbf{f}(\bar{\mathbf{z}} + h\mathbf{e}_i) - \mathbf{f}(\bar{\mathbf{z}} - h\mathbf{e}_i)) \\ &\quad \times (\mathbf{f}(\bar{\mathbf{z}} + h\mathbf{e}_i) - \mathbf{f}(\bar{\mathbf{z}} - h\mathbf{e}_i))^T \\ &= \frac{1}{4h^2} \sum_{i=1}^n (\mathbf{f}(\bar{\mathbf{x}} + h\mathbf{s}_{x,i}) - \mathbf{f}(\bar{\mathbf{x}} - h\mathbf{s}_{x,i})) \\ &\quad \times (\mathbf{f}(\bar{\mathbf{x}} + h\mathbf{s}_{x,i}) - \mathbf{f}(\bar{\mathbf{x}} - h\mathbf{s}_{x,i}))^T, \\ \mathbf{P}_{xy} &= \mathbf{E}[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})^T] \\ &= \mathbf{E}[\mathbf{S}_x \Delta \mathbf{z} (\tilde{\mathbf{D}}_{\Delta z} \tilde{\mathbf{f}})^T] \\ &= \frac{1}{2h} \sum_{i=1}^n \mathbf{s}_{x,i} (\mathbf{f}(\bar{\mathbf{x}} + h\mathbf{s}_{x,i}) - \mathbf{f}(\bar{\mathbf{x}} - h\mathbf{s}_{x,i}))^T,\end{aligned}\tag{8}$$

where  $\mathbf{s}_{x,i}$  is the  $i$ th column of the matrix  $\mathbf{S}_x$ .

Consider the following nonlinear dynamic system with states to be estimated:

$$\begin{aligned}\mathbf{x}_{k+1} &= \mathbf{f}(\mathbf{x}_k) + \boldsymbol{\omega}_k \\ \mathbf{y}_k &= \mathbf{h}(\mathbf{x}_k) + \boldsymbol{\nu}_k,\end{aligned}\tag{9}$$

where  $\boldsymbol{\omega}_k$  and  $\boldsymbol{\nu}_k$  are assumed to be independent and identically distributed and independent of current and past states, such that  $\boldsymbol{\omega}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{Q}_k)$  and  $\boldsymbol{\nu}_k \sim \mathbf{N}(\mathbf{0}, \mathbf{R}_k)$ .

The DDF takes the same predictor-corrector structures in the EKF and can be described as follows.

*Step 1* (initialization). Suppose the state distribution at  $k$  instant is  $\mathbf{x}_k \sim \mathbf{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k)$ , where  $\hat{\mathbf{x}}_k$  and  $\mathbf{P}_k$  are obtained by

$$\begin{aligned}\hat{\mathbf{x}}_k &= \mathbf{E}[\mathbf{x}_k], \\ \mathbf{P}_k &= \mathbf{E}[(\mathbf{x} - \hat{\mathbf{x}}_k)(\mathbf{x} - \hat{\mathbf{x}}_k)^T].\end{aligned}\tag{10}$$

*Step 2* (square Cholesky factorizations). Consider the following:

$$\begin{aligned}\mathbf{P}_k &= \mathbf{S}_x \mathbf{S}_x^T \\ \mathbf{S}_{x\hat{\mathbf{x}}} &= \frac{1}{2h} \left\{ \mathbf{f}_i(\hat{\mathbf{x}}_k + h\mathbf{s}_{x,j}) - \mathbf{f}_i(\hat{\mathbf{x}}_k - h\mathbf{s}_{x,j}) \right\},\end{aligned}\tag{11}$$

where  $\mathbf{s}_{x,i}$  is the  $j$ th column of  $\mathbf{S}_x$ .

*Step 3* (state and covariance propagation). One has

$$\hat{\mathbf{x}}_{k+1}^- = \mathbf{f}(\hat{\mathbf{x}}_k) \tag{12}$$

$$\begin{aligned}\mathbf{P}_{k+1}^- &= \mathbf{S}_{x\hat{\mathbf{x}}}(\mathbf{S}_{x\hat{\mathbf{x}}})^T + \mathbf{Q}_k \\ &= \mathbf{S}_x^-(\mathbf{S}_x^-)^T\end{aligned}\tag{13}$$

$$\mathbf{S}_{y\hat{\mathbf{x}}} = \frac{1}{2h} \left\{ \mathbf{h}_i(\hat{\mathbf{x}}_{k+1}^- + \mathbf{s}_{x,j}^-) - \mathbf{h}_i(\hat{\mathbf{x}}_{k+1}^- - \mathbf{s}_{x,j}^-) \right\}, \tag{14}$$

where  $\hat{\mathbf{x}}_{k+1}^-$  is the predicted state and  $\mathbf{P}_{k+1}^-$  is the predicted covariance matrix.

*Step 4* (observation and innovation covariance propagation). Consider

$$\hat{\mathbf{y}}_{k+1}^- = \mathbf{h}(\hat{\mathbf{x}}_{k+1}^-) \tag{15}$$

$$\mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}} = (\mathbf{S}_{y\hat{\mathbf{x}}})(\mathbf{S}_{y\hat{\mathbf{x}}})^T + \mathbf{R}_k \tag{16}$$

$$\mathbf{P}_{k+1}^{\mathcal{X}\mathcal{Y}} = \mathbf{S}_x^-(\mathbf{S}_{y\hat{\mathbf{x}}})^T, \tag{17}$$

where  $\hat{\mathbf{y}}_{k+1}^-$  is the predicted observation vector,  $\mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}}$  is the innovation covariance matrix, and  $\mathbf{P}_{k+1}^{\mathcal{X}\mathcal{Y}}$  is the cross correlation matrix.

*Step 5* (update). Consider the following:

$$\begin{aligned}\boldsymbol{\kappa}_{k+1} &= \mathbf{P}_{k+1}^{\mathcal{X}\mathcal{Y}} (\mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}})^{-1} \\ \mathbf{P}_{k+1}^+ &= \mathbf{P}_{k+1}^- - \boldsymbol{\kappa}_{k+1} \mathbf{P}_{k+1}^{\mathcal{Y}\mathcal{Y}} \boldsymbol{\kappa}_{k+1}^T \\ \mathbf{v}_{k+1} &= \mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1}^- \\ \hat{\mathbf{x}}_{k+1}^+ &= \hat{\mathbf{x}}_{k+1}^- + \boldsymbol{\kappa}_{k+1} \mathbf{v}_{k+1},\end{aligned}\tag{18}$$

where  $\boldsymbol{\kappa}_{k+1}$  is the gain,  $\mathbf{P}_{k+1}^+$  is the updated covariance matrix,  $\mathbf{v}_{k+1}$  is the innovation vector, and  $\hat{\mathbf{x}}_{k+1}^+$  is the updated estimated state.

*2.2. Adaptive DDF Algorithm.* As stated previously, the DDF algorithm assumes a complete prior knowledge of the process and the measurement noise statistics  $\mathbf{Q}_k$  and  $\mathbf{R}_k$ . However,  $\mathbf{Q}_k$  and  $\mathbf{R}_k$  are unknown in most applications, and incorrect priori noise statistics can lead to performance degradation or even divergence for the solution. One of the effective ways to overcome this weakness is to use an algorithm to adapt the noise statistics. In this paper, we propose using a support vector regression adaptive scheme of the DDF to adjust  $\mathbf{Q}_k$  and  $\mathbf{R}_k$ , respectively.

*2.2.1. Support Vector Regression.* The principles underlying support vector regression (SVR) developed by Vapnik and are presented in several works (see [26, 27]).

Given the train set  $TS = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_m, y_m)\} \in (\mathbf{R}^n \times \mathbf{R})^m$ , where  $\mathbf{x}_i \in \mathbf{R}^n$ ,  $y_i \in \mathbf{R}$ ,  $i = 1, 2, \dots, m$ , the SVR problem can be defined as solving for the nonlinear function  $g(\mathbf{x})$  about  $\mathbf{x} \in \mathbf{R}^n$  to construct a relationship between the output and an arbitrary input  $\mathbf{x}$ :

$$y(\mathbf{x}, \boldsymbol{\omega}) = \sum_{i=1}^m \omega_i g(\mathbf{x}_i) + b = (\boldsymbol{\omega} \cdot \mathbf{g}(\mathbf{x})) + b. \quad (19)$$

In [26], the author shows that changing the regression estimate minimizes the risk functional by using the following form:

$$\begin{aligned} y(\mathbf{x}, \boldsymbol{\alpha}) &= \sum_{i=1}^m (\alpha_i^* - \alpha_i) (g(\mathbf{x}_i) \cdot g(\mathbf{x})) + b \\ &= \sum_{i=1}^m (\alpha_i^* - \alpha_i) K(\mathbf{x}_i, \mathbf{x}) + b, \end{aligned} \quad (20)$$

where  $\alpha_i^*$  and  $\alpha_i$  are Lagrange multipliers that satisfy the condition  $\alpha_i, \alpha_i^* \geq 0$ ,  $\alpha_i \alpha_i^* = 0$  and  $K(\mathbf{x}_i, \mathbf{x})$  is a kernel function that satisfies Mercer's condition. In this paper, we use a translation invariant Gaussian kernel, that is,  $= \exp(-\|\mathbf{x}_i - \mathbf{x}\|^2 / 2\sigma^2)$ .

**2.2.2. Adaptive Scheme Based on SVR ( $\mathbf{Q}_k$  Is Fixed).** The covariance matrix  $\mathbf{R}_k$  represents the accuracy of the measurement instrument. If we assume that the noise covariance  $\mathbf{Q}_k$  is completely known, then we can derive the SVR algorithm to estimate the measurement noise covariance  $\mathbf{R}_k$  by defining an adaptive factor  $\Delta r_k$  to get the form:

$$\mathbf{R}_k = \Delta r_k \mathbf{R}, \quad (21)$$

where  $\mathbf{R}$  is the constant noise covariance matrix.

This work uses such an SVR algorithm to derive the adaptive factor at time instant  $k$ , so as to estimate the value of  $\mathbf{R}_k$  during the algorithm's execution.

The innovation sequence  $\mathbf{v}_{k+1}$  has a theoretical covariance  $\mathbf{P}_{k+1}^{\text{yy}}$ , as (16) shows. The actual residual covariance  $\mathbf{P}_{k+1}^{\text{yy}}$  can be approximated using its sample covariance by averaging inside a moving window of size  $N$ , where

$$\mathbf{C}_k = \frac{1}{N} \sum_{i=k-N+1}^k \mathbf{v}_{k+1} \mathbf{v}_{k+1}^T. \quad (22)$$

If  $\mathbf{C}_k$  differs from  $\mathbf{P}_{k+1}^{\text{yy}}$ , then the diagonal elements of  $\mathbf{R}_k$  can be adjusted to minimize the difference as much as possible. The size of this difference is given by

$$\mathbf{DOM}_k = \text{diag}(\mathbf{P}_{k+1}^{\text{yy}} - \mathbf{C}_k) \in \mathbf{R}^n, \quad (23)$$

where the function  $\text{diag}$  denotes the diagonal elements of the matrix. If the elements of  $\mathbf{DOM}_k$  are rounded with zero, the covariance matrix  $\mathbf{R}_k$  is likely accurate; otherwise there may be a large deviation between values.

To adjust  $\mathbf{R}_k$ , we define the SVR train set  $TS_{\Delta r}$  as  $TS_{\Delta r} = \{(\mathbf{DOM}_{k-N+1}, \Delta r_{k-N+1}), \dots, (\mathbf{DOM}_k, \Delta r_k)\}$ . With this form, we can then create a corresponding function about the

adaptive factor  $\Delta r_k$  and  $\mathbf{DOM}_k$  based on (19), such that when  $\mathbf{DOM}_{k+1}$  comes, then  $\Delta r_{k+1}$  can be resolved by the function.

To avoid having any element within the train set be too large to affect the accuracy of  $\Delta r_k$ , we use a normalized function of the train set  $TS$ ; that is,  $TS'_{\Delta r} = \{(\mathbf{DOM}'_{k-N+1}, \Delta r'_{k-N+1}), \dots, (\mathbf{DOM}'_k, \Delta r'_k)\}$ , where  $\mathbf{DOM}'_k$  and  $r'_k$  can be expressed as follows:

$$\begin{aligned} \mathbf{DOM}'_k &= \frac{1}{N} \left\{ \sum_{i=1}^N \mathbf{DOM}_{k-i+1, j} \right\} \quad j = 1, 2, \dots, n, \\ \Delta r'_k &= \frac{1}{N} \sum_{i=1}^N \Delta r_{k-i+1}. \end{aligned} \quad (24)$$

**2.2.3. Adaptive Scheme Based on SVR ( $\mathbf{R}_k$  Is Fixed).** Assuming that the noise covariance  $\mathbf{R}_k$  is completely known, we can derive an SVR algorithm to estimate the measurement noise covariance  $\mathbf{Q}_k$ . From (13), (14), and (16), we can deduce that a change in  $\mathbf{Q}_k$  will affect the covariance matrix  $\mathbf{P}_{k+1}^{\text{yy}}$ ; if we increase  $\mathbf{Q}_k$ , then  $\mathbf{P}_{k+1}^{\text{yy}}$  also increases. We can adjust  $\mathbf{Q}_k$  in the SVR by deliberately mismatching  $\mathbf{P}_{k+1}^{\text{yy}}$  and  $\mathbf{C}_{k+1}$ .

We first define an adaptive factor  $\Delta r_k$ , where  $\mathbf{R}_k$  has the following form:

$$\mathbf{Q}_k = \Delta q_k \mathbf{Q}, \quad (25)$$

where  $\mathbf{Q}$  is the constant noise covariance matrix.

We can then define the SVR train set  $TS_{\Delta q}$  as  $TS_{\Delta q} = \{(\mathbf{DOM}_{k-N+1}, \Delta r_{k-N+1}), \dots, (\mathbf{DOM}_k, \Delta r_k)\}$ . At this point, the solution process is the same as for solving  $\Delta r_k$ .

### 3. Monte Carlo Simulation Results and Discussion

In this section, we report the experimental results obtained by applying SVRADDF to the nonlinear state estimation of a nonmaneuvering target in an underwater tracking control scenario and a maneuvering target in an air-traffic control scenario. To demonstrate the performance of the SVRADDF algorithm, we compare its performance against a DDF algorithm.

**3.1. Underwater Nonmaneuvering Target Bearing-Only Tracking Control Scenario.** We consider a bearing-only tracking control scenario, where an underwater target executes a uniform motion in a horizontal plane but unknown velocity, while a passive sonar platform performs a uniform circular motion in a horizontal plane. Figure 2 shows a representative trajectory of the target and the passive sonar platform. The kinematics of the relative motion between the target and the platform can be modeled using the following linear process equation:

$$\mathbf{x}_{k+1} = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_k + \mathbf{w}_k. \quad (26)$$



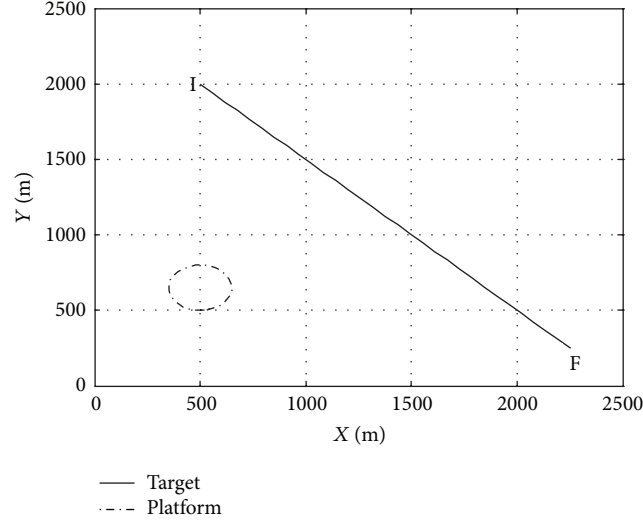


FIGURE 2: Target trajectory (I-initial position and F-final position) and sonar trajectory.

Here, the state of the equation is  $\mathbf{x} = [x \ y \ \dot{x} \ \dot{y}]$ , where  $x$  and  $y$  denote position and  $\dot{x}$  and  $\dot{y}$  denote velocity in the  $x$  and  $y$  directions, respectively, and  $T$  is the time interval between two consecutive measurements and the process noise  $\mathbf{w}_k \sim N(0, \mathbf{Q})$  with a nonsingular covariance where

$$\mathbf{Q} = q_1 \times \begin{bmatrix} \frac{T^2}{2} \\ \frac{T^2}{2} \\ T \\ T \end{bmatrix}. \quad (27)$$

In (27), the parameter  $q_1$  is related to process noise intensities. The measurement equation is written as follows:

$$\mathbf{y}_k = \theta_k = \tan^{-1} \left( \frac{y_k}{x_k} \right) + v_k, \quad (28)$$

where the measurement noise  $v_k \sim N(0, R)$  with a nonsingular covariance.

Given the following initial conditions:

$$\begin{aligned} T &= 1 \text{ s} \\ q_1 &= 0.0001 \text{ m}^2 \text{ s}^{-3} \\ R &= 0.02 \text{ mrad}, \end{aligned} \quad (29)$$

the true initial state is

$$\mathbf{x}_0 = [0 \text{ m} \ 1500 \text{ m} \ 0 \text{ m s}^{-1} \ 0 \text{ m s}^{-1}]^T \quad (30)$$

and the associated covariance is

$$\mathbf{P}_0 = \text{diag} [100 \text{ m}^2 \ 2000 \text{ m}^2 \ 1 \text{ m}^2 \text{ s}^{-2} \ 1 \text{ m}^2 \text{ s}^{-2}]. \quad (31)$$

The initial estimate state  $\hat{\mathbf{x}}_0$  was chosen randomly from  $N(\mathbf{x}_0, \mathbf{P}_0)$  in each run, and the total number of scans per run was 1000.

To provide a fair comparison, we performed 50 independent Monte Carlo runs. To track the underwater target, we used both the SVRADDf and the DDF algorithms and compared their performance. The adaptive factor was set to  $\Delta q_k = \Delta r_k = 0.1$ . Both of the filters were initialized with the same initial conditions for each run.

Performance metrics: to compare the nonlinear performance of the filters, we used the root mean square error (RMSE) of the target position and velocity. The RMSE yields a combined measure of the bias and variance of a filter estimate. The RMSE of the position at time  $k$  was found using

$$\text{RMSE}_{\text{pos}}(k) = \sqrt{\frac{1}{N} \sum_{i=1}^N ((x_k^i - \hat{x}_k^i)^2 + (y_k^i - \hat{y}_k^i)^2)}, \quad (32)$$

where  $(x_k^i, y_k^i)$  and  $(\hat{x}_k^i, \hat{y}_k^i)$  are the true and estimated positions, respectively, in the  $i$ th Monte Carlo run. The form for the RMSE of the velocity is similar.

Figures 3 and 4 show the estimated RMSE in target position and velocity. The SVRADDf uses SVR to adjust the adaptive factor during algorithm execution, which leads to a marginally better performance compared to the DDF, as seen in the figures.

**3.2. Maneuvering Target Tracking in the Air-Traffic Control Scenario.** A typical air-traffic control scenario was considered next, where an aircraft executes a maneuvering turn in a horizontal plane at a constant and known turn rate  $\Omega$ . Figure 5 shows a representative trajectory of the aircraft.

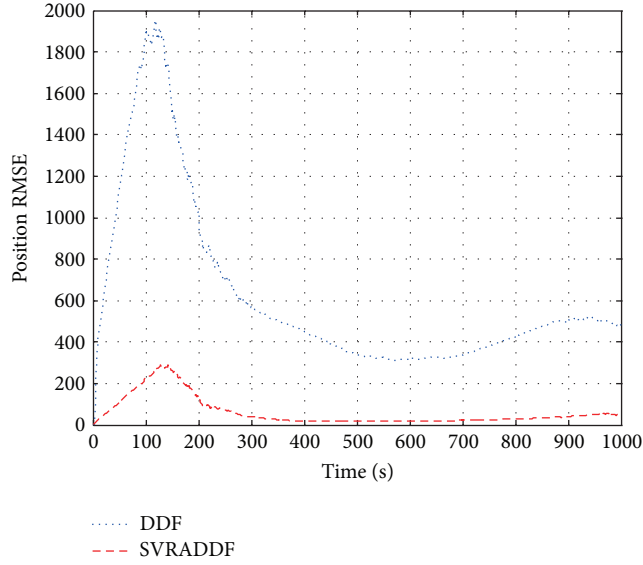


FIGURE 3: RMSE in position for DDF and SVRADDF.

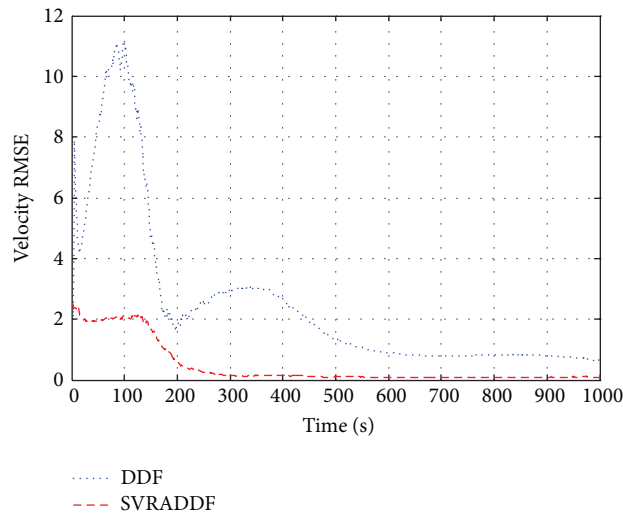


FIGURE 4: RMSE in velocity for DDF and SVRADDF.

The kinematics of the turning motion can be modeled using the following nonlinear process equation:

$$\mathbf{x}_k = \begin{bmatrix} 1 & \frac{\sin \Omega T}{\Omega} & 0 & -\frac{1 - \cos \Omega T}{\Omega} \\ 0 & \cos \Omega T & 0 & -\sin \Omega T \\ 0 & \frac{1 - \cos \Omega T}{\Omega} & 1 & \frac{\sin \Omega T}{\Omega} \\ 0 & \sin \Omega T & 0 & \cos \Omega T \end{bmatrix} \times \mathbf{x}_{k-1} + \mathbf{w}_{k-1}. \quad (33)$$

The state of the aircraft is given by  $\mathbf{x} = [x \ \dot{x} \ y \ \dot{y}]^T$ , where  $x$  and  $y$  denote position and  $\dot{x}$  and  $\dot{y}$  denote velocity in the  $x$  and  $y$  directions, respectively, and  $T$  is the time

interval between two consecutive measurements and the process noise  $\mathbf{w}_k \sim N(0, \mathbf{Q})$  with a nonsingular covariance, where

$$\mathbf{Q} = q_1 \times \begin{bmatrix} \frac{T^3}{3} & \frac{T^2}{2} & 0 & 0 \\ \frac{T^2}{2} & T & 0 & 0 \\ 0 & 0 & \frac{T^3}{3} & \frac{T^2}{2} \\ 0 & 0 & \frac{T^2}{2} & T \end{bmatrix}. \quad (34)$$

The parameter  $q_1$  related to process noise intensities. A passive radar is fixed at the origin and equipped to measure

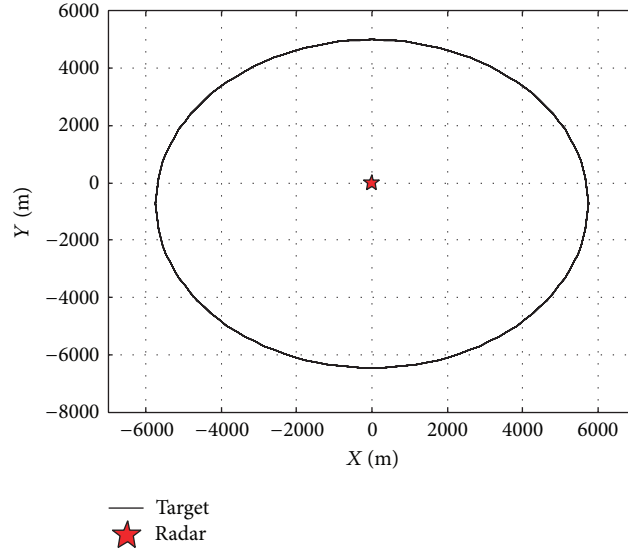


FIGURE 5: Aircraft trajectory and radar location.

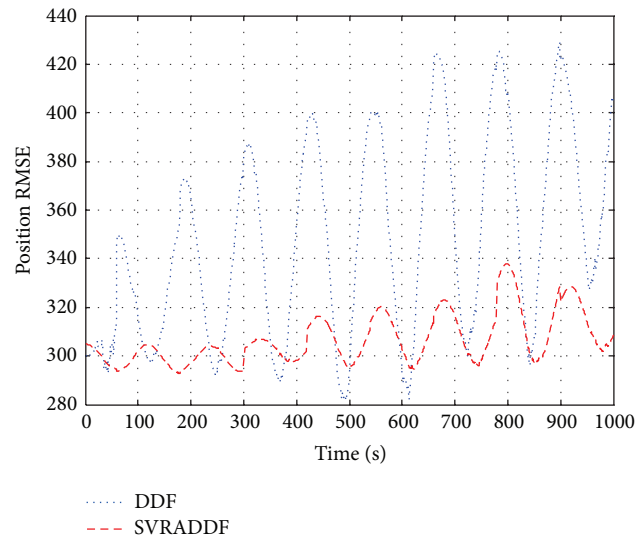


FIGURE 6: RMSE in position for DDF and SVRADDF.

the bearing  $\theta$ . The measurement equation is written as follows:

$$y_k = \theta_k = \tan^{-1} \left( \frac{y_k}{x_k} \right) + v_k, \quad (35)$$

where the measurement noise  $v_k \sim N(0, R)$ .

The parameters used in this simulation were  $T = 1\text{ s}$ ,  $\Omega = -3^\circ\text{ s}^{-1}$ ,  $q_1 = 0.1\text{ m}^2\text{ s}^{-3}$ , and  $R = \sqrt{10}\text{ mard}$ . The true initial state of the aircraft was  $\mathbf{x}_0 = [1000\text{ m } 300\text{ m s}^{-1} 1000\text{ m } 0\text{ m s}^{-1}]^T$  and the associated covariance matrix was  $\mathbf{P}_0 = \text{diag}[100\text{ m}^2 10\text{ m}^2\text{ s}^{-2} 100\text{ m}^2 10\text{ m}^2\text{ s}^{-2}]$ . The initial state  $\hat{\mathbf{x}}_0$  for the filters was chosen randomly from  $N(\mathbf{x}_0, \mathbf{P}_0)$  in each Monte Carlo run, and the simulation time per run was 1000.

For a fair comparison, we performed 100 independent Monte Carlo runs for each filter. To track the maneuvering aircraft, we used both the SVRADDF and the DDF algorithms and compared their performance. The adaptive factor was set to  $\Delta q_k = 0.3$  and  $\Delta r_k = 0.5$ . Both filters were initialized with the same initial conditions for each run.

Figures 6 and 7 show the estimate RMSE in target position and velocity for the SVRADDF and DDF filters. Not surprisingly, both filters exhibit divergence due to a mismatch between the initial filter design assumption and the Gaussian noise nature of the problem. The SVRADDF filter exhibits marginally better performance compared to the DDF since it was able to adjust the statistical properties of the noise during execution.



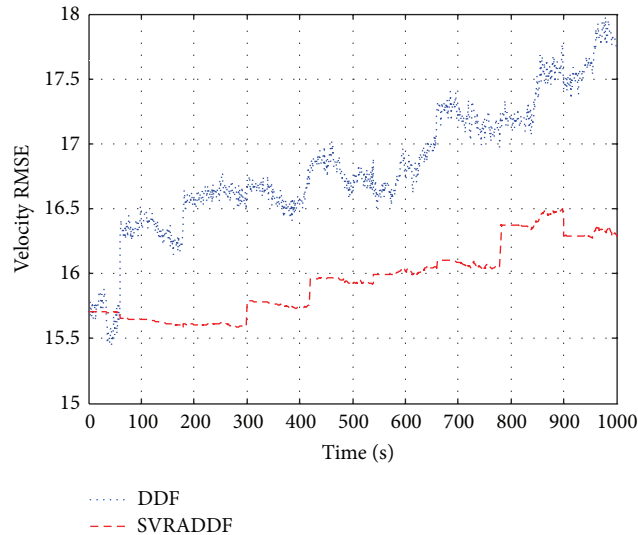


FIGURE 7: RMSE in velocity for DDF and SVRADDF.

#### 4. Conclusions

This work has proposed and developed an innovation-based SVRADDF algorithm. The algorithm introduces an adaptive factor estimated using SVR, which allows for estimation of the noise statistical characteristics of nonlinear stochastic systems. The SVRADDF algorithm avoids instability and divergence in the solution which is caused by incorrect statistical characteristics of the noise. Monte Carlo simulation results of an underwater nonmaneuvering bearing-only target tracking system and a maneuvering target bearing-only tracking system in an air-traffic control setting showed that the SVRADDF algorithm provides better state estimation accuracy than a traditional DDF algorithm.

Although the SVRADDF algorithm has showed better performance under Monte Carlo simulation, there are still several challenging issues to be considered for future study. To improve its feasibility and effectiveness in the complex environment, more comprehensive and detailed studies are still needed to solve under the nonlinear and non-Gaussian noise conditions. Since the algorithm has only been tested under Monte Carlo simulation, the following work might be testing the algorithm under trial data to approve its better performance online.

#### Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

#### Acknowledgments

This research work is supported by the National Natural Science Foundation of China (Grant no. E091002/50979017), Ph.D. Programs Foundation of Ministry of Education of China and Basic Technology (Grant no. 20092304110008), and Harbin Science and Technology Innovation Talents of

Special Fund Project (Outstanding Subject Leaders) (no. 2012RFXXG083).

#### References

- [1] Y. Bar-Shalom and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*, John Wiley & Sons, New York, NY, USA, 2001.
- [2] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *Proceedings of the American Control Conference*, pp. 1628–1632, June 1995.
- [3] B. M. Bell and F. W. Cathey, "Iterated Kalman filter update as a Gauss-Newton method," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 294–297, 1993.
- [4] K.-H. Kim, G.-I. Jee, C.-G. Park, and J.-G. Lee, "The stability analysis of the adaptive fading extended kalman filter using the innovation covariance," *International Journal of Control, Automation and Systems*, vol. 7, no. 1, pp. 49–56, 2009.
- [5] X.-X. Wang, L. Zhao, Q.-X. Xia, W. Cao, and L. Li, "Design of unscented Kalman filter with correlative noises," *Control Theory & Applications*, vol. 27, no. 10, pp. 1362–1368, 2010.
- [6] X.-X. Wang, L. Zhao, Q. Pan, Q.-X. Xia, and W. Hong, "Design of UKF with correlative noises based on minimum mean square error estimation," *Control and Decision*, vol. 25, no. 9, pp. 1393–1398, 2010.
- [7] X.-X. Wang, Y. Liang, Q. Pan, C.-H. Zhao, and H.-Z. Li, "Unscented Kalman filter for nonlinear systems with colored measurement noise," *Acta Automatica Sinica*, vol. 38, no. 6, pp. 986–998, 2012.
- [8] W. F. Leven and A. D. Lanterman, "Unscented Kalman filters for multiple target tracking with symmetric measurement equations," *IEEE Transactions on Automatic Control*, vol. 54, no. 2, pp. 370–375, 2009.
- [9] J.-S. Kim, D.-R. Shin, and E. Serpedin, "Adaptive multiuser receiver with joint channel and time delay estimation of CDMA signals based on the square-root unscented filter," *Digital Signal Processing*, vol. 19, no. 3, pp. 504–520, 2009.
- [10] W. Fan and Y. Li, "Accuracy analysis of sigma-point Kalman filters," in *Proceedings of the Chinese Control and Decision*

- Conference (CCDC '09)*, pp. 2883–2888, Guilin, China, June 2009.
- [11] T. Lefebvre, H. Bruyninckx, and J. de Schutter, “Kalman filters for non-linear systems: a comparison of performance,” *International Journal of Control*, vol. 77, no. 7, pp. 639–653, 2004.
  - [12] J. Dunik and M. Simandl, “Performance analysis of derivative-free filters,” in *Proceedings of the 44th IEEE Conference on Decision and Control*, pp. 1941–1946, Seville, Spain, 2005.
  - [13] M. Nørgaard, N. K. Poulsen, and O. Ravn, “New developments in state estimation for nonlinear systems,” *Automatica*, vol. 36, no. 11, pp. 1627–1638, 2000.
  - [14] L. Deok-Jin, *Nonlinear Bayesian filtering with applications to estimation and navigation [Ph.D. thesis]*, 2005.
  - [15] C. Wang, J. Zhang, and J. Mu, “Maximum likelihood-based iterated divided difference filter for nonlinear systems from discrete noisy measurements,” *Sensors*, vol. 12, no. 7, pp. 8912–8929, 2012.
  - [16] N. Subrahmanya and Y. C. Shin, “Adaptive divided difference filtering for simultaneous state and parameter estimation,” *Automatica*, vol. 45, no. 7, pp. 1686–1693, 2009.
  - [17] X. Luo, I. Hoteit, and I. M. Moroz, “On a nonlinear Kalman filter with simplified divided difference approximation,” *Physica D: Nonlinear Phenomena*, vol. 241, no. 6, pp. 671–680, 2012.
  - [18] Z. Ali, M. A. Landolsi, and M. Deriche, “Multiuser parameter estimation using divided difference filter in CDMA systems,” *International Journal of Advanced Science and Technology*, vol. 5, pp. 35–50, 2009.
  - [19] P. G. Bhale, P. N. Dwivedi, P. Kumar, and A. Bhattacharyaa, “Estimation of ballistic coefficient of reentry vehicle with divided difference filtering using noisy RF seeker data,” in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT '06)*, pp. 1087–1092, Mumbai, India, December 2006.
  - [20] M. Jing, W. Chang-yuan, Z. Juan, and C. Fang, “State estimation of reentry target based on divided difference filter,” *Journal of Xi'an Technological University*, vol. 31, no. 4, pp. 377–381, 2011.
  - [21] P. Tripathy, S. C. Srivastava, and S. N. Singh, “A divide-by-difference-filter based algorithm for estimation of generator rotor angle utilizing synchrophasor measurements,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 6, pp. 1562–1570, 2010.
  - [22] M. Ahmadi, A. Khayatian, and P. Karimaghvae, “Attitude estimation by divided difference filter in quaternion space,” *Acta Astronautica*, vol. 75, pp. 95–107, 2012.
  - [23] P. Setoodeh, A. Khayatian, and E. Farjah, “Attitude estimation by divided difference filter-based sensor fusion,” *The Journal of Navigation*, vol. 60, no. 1, pp. 119–128, 2007.
  - [24] C. Wu and C. Han, “Second-order divided difference filter with application to ballistic target tracking,” in *Proceedings of the 7th World Congress on Intelligent Control and Automation (WCICA '08)*, pp. 6336–6341, Chongqing, China, June 2008.
  - [25] D.-J. Jwo, M.-Y. Hsieh, and S.-Y. Lai, “GPS navigation processing using the quaternion-based divided difference filter,” *GPS Solutions*, vol. 14, no. 3, pp. 217–228, 2010.
  - [26] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, Berlin, Germany, 1995.
  - [27] L. Guozheng, W. Meng, and Z. Huajun, *An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods*, Publishing House of Electronics Industry, Beijing, China, 2004.