International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol.9, No.9, (2016), pp.167-176
http://dx.doi.org/10.14257/ijsip.2016.9.9.16

Maximum Likelihood Principle Based Adaptive UKF Algorithm

Li Guo[1,2*] and De-gen Huang[2]

[1]School of Information Science and Engineering, Dalian University of Technology, Dalian, Liaoning 116024 China
[2]School of Information Science and Engineering, Dalian Polytechnic University, Dalian, Liaoning 116034 China
[1]guoli1117@163.com

Abstract

In this paper, we investigate the state estimation problem of nonlinear systems under the condition that the prior statistical characteristic of noise is unknown. An adaptive unscented Kalman filter (UKF) is proposed. In this algorithm, the maximum likelihood principle is applied to establish the log likelihood function with the unknown noise statistical characteristics. Then, the noise property estimation problem is transformed into the maximization of the mean of the log likelihood function, which can be achieved by using the expectation maximization algorithm. Finally, a suboptimal adaptive UKF can be obtained. Simulations show that the proposed adaptive UKF algorithm can deal with the problem of filtering accuracy declination of the traditional UKF when the prior noise statistical characteristic is unknown. The proposed algorithm can estimate the statistical parameters online.

Keywords: *Nonlinear filtering; adaptive UKF algorithm; noise statistics estimator, maximum likelihood principle*

1. Introduction

The problem of nonlinear filtering widely exists in the practical applications. The key point of nonlinear filtering is the calculation of the posteriori distribution of the random vector along the nonlinear functions. To deal with this problem, numerous methods have been proposed. The extended Kalman filter (EKF) has been widely used in recent years [1]. By using the first order linearization, the nonlinear system is approximate as a linear one, and thus, the standard Kalman filter can be applied to estimate the states of the nonlinear system. The accuracy of the linearization and the requirement of the Jacobin matrix lead the limitation of the usage of EKF in some practical applications. Particle filter (PF) is a sequential importance sampling (SIS) algorithm based on Bayesian estimation [2]. It has been widely used to deal with the state estimation problem for strong nonlinear system.

Unscented Kalman filter (UKF) based on unscented transformation, which is originally proposed by Julier [3], is a widely used nonlinear filtering algorithm. The unscented transformation can approximate the nonlinear posteriori distribution with the accuracy of at least more than second order. Meanwhile, it no longer needs to calculate the Jacobin matrix and the volume of calculation is less than PF algorithm.

However, when the prior noise statistics of the noise is unknown, the accuracy of UKF will decrease or the system will even diverge [4]. This is because the theoretical basis of UKF algorithm is the variance minimization principle. The premise is that the prior noise statistics should be known accurately. In order to solve this problem, several adaptive algorithms have been proposed to estimate the

*Corresponding Author

noise information online, such as Bayesian algorithm [5-6], maximum likelihood algorithm [7-8], correlation algorithm [9-10] and covariance matching algorithm [11-12], *etc.*

The Bayesian algorithm usually involves multiple operations. Meanwhile, the calculation amount of this algorithm is large, and the closed solution usually cannot be obtained. Hence the application of this algorithm is limited. The steady-state estimation error of covariance matching algorithm cannot be eliminated, which decrease the accuracy of this algorithm. Correlation algorithm can only be used in linear system. The maximum likelihood algorithm has been widely investigated since it can establish the probability density function with statistics parameters directly. Meanwhile, the calculation amount of this algorithm is moderate for practical application.

From the analysis above, the maximum likelihood based adaptive UKF algorithm is proposed in this paper. The maximum likelihood principle is applied to establish the Logarithmic likelihood function, based on which the noise statistics can be introduced. Then, the Expectation Maximization (EM) algorithm is employed to simplify the second-order moment calculation of the noise. The proposed strategy can estimate the system states without the prior information of noise statistics.

The following of this paper is organized as follows. In Section II, the state estimation problem is formulated and traditional UKF algorithm is introduced. In Section III, the noise statistics estimator is proposed based on maximum likelihood principle and EM algorithm. Simulations results in Section IV show the effectiveness of the proposed algorithm, followed by Conclusions in Section V.

2. Problem Formulation and Traditional UKF

Consider the discrete-time nonlinear system with additive Gaussian noise:

$$\begin{cases} \boldsymbol{x}_k = \boldsymbol{f}\left(\boldsymbol{x}_{k-1}\right) + \boldsymbol{w}_{k-1} \\ \boldsymbol{z}_k = \boldsymbol{h}\left(\boldsymbol{x}_k\right) + \boldsymbol{v}_k \end{cases} \tag{1}$$

where $k \ge 0$ is discrete-time variable, $\boldsymbol{x} \in \mathbf{R}^n$ is state vector, $\boldsymbol{z} \in \mathbf{R}^m$ is output vector. The nonlinear function $\boldsymbol{f} \in \mathbf{R}^n \to \mathbf{R}^n$, $\boldsymbol{h} \in \mathbf{R}^n \to \mathbf{R}^m$. The process noise $\boldsymbol{w}_{k-1}$ and measurement noise $\boldsymbol{v}_k$ are n'th and m'th order Gaussian noise, which satisfy:

$$\begin{cases} \mathrm{E}\left[\boldsymbol{w}_k\right] = \boldsymbol{q}, \ \ \mathrm{cov}\left(\boldsymbol{w}_k, \boldsymbol{w}_j\right) = \boldsymbol{Q} \cdot \delta_{k,j} \\ \mathrm{E}\left[\boldsymbol{v}_k\right] = \boldsymbol{r}, \ \ \mathrm{cov}\left(\boldsymbol{v}_k, \boldsymbol{v}_j\right) = \boldsymbol{R} \cdot \delta_{k,j} \\ \mathrm{cov}\left(\boldsymbol{w}_k, \boldsymbol{v}_j\right) = 0 \end{cases} \tag{2}$$

The steps of traditional UKF algorithm are given as follows:

STEP1. System states initiation.

STEP2. Time update.

From the given sampling strategy, the Sigma point $\left\{\boldsymbol{\chi}_{i,k-1}\right\}, i = 0, \cdots, 2n$ that with the mean of $\hat{\boldsymbol{x}}_{k-1}$ and covariance of $\boldsymbol{P}_{k-1}$ can be obtained as:

$$\begin{cases} \boldsymbol{\chi}_{0,k-1} = \hat{\boldsymbol{x}}_{k-1}, \\ \boldsymbol{\chi}_{i,k-1} = \hat{\boldsymbol{x}}_{k-1} + \left(\sqrt{\left(n+\lambda\right)\boldsymbol{P}_{k-1}}\right)_i, i = 1, 2, \cdots n, \\ \boldsymbol{\chi}_{i,k-1} = \hat{\boldsymbol{x}}_{k-1} - \left(\sqrt{\left(n+\lambda\right)\boldsymbol{P}_{k-1}}\right)_i, i = n+1, n+2, \cdots 2n; \end{cases} \tag{3}$$

The weights of corresponding Sigma points are given as:

$$\begin{cases} \omega_0^{(m)} = \lambda/(n+\lambda), \\ \omega_0^{(c)} = \lambda/(n+\lambda) + (1-\alpha^2+\beta), \\ \omega_i^{(m)} = \omega_i^{(c)} = 0.5/(n+\lambda), i = 1,2,\cdots 2n; \end{cases} \tag{4}$$

where operator $\sqrt{\cdot}$ represents Cholesky decomposition of a matrix. α is the extendible extent of Sigma near the mean, β is the adjusting parameter, $\lambda = \alpha^2(n+k) - n$. Here we choose: $\alpha = 1$, $\beta = 0$, $k = 2$.

The transmitting effect of the Sigma point along the nonlinear function is given as:

$$\boldsymbol{\chi}_{i,k|k-1} = \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right) + \boldsymbol{q} \quad i = 0,\cdots,2n \tag{5}$$

Then we get the posterior mean and covariance as:

$$\hat{\boldsymbol{x}}_{k|k-1} = \sum_{i=0}^{2n} \omega_i^{(m)} \cdot \boldsymbol{\chi}_{i,k|k-1} = \sum_{i=0}^{2n} \omega_i^{(m)} \cdot \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right) + \boldsymbol{q} \tag{6}$$

$$\boldsymbol{P}_{k|k-1} = \sum_{i=0}^{2n} \omega_i^{(c)} \cdot \left(\boldsymbol{\chi}_{i,k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}\right) \cdot \left(\boldsymbol{\chi}_{i,k|k-1} - \hat{\boldsymbol{x}}_{k|k-1}\right)^{\mathrm{T}} + \boldsymbol{Q} \tag{7}$$

STEP3. Measurement Update.

From the given sampling strategy, the Sigma point $\{\boldsymbol{\chi}_{i,k|k-1}\}, i = 0,\cdots,2n$ that with the mean of $\hat{\boldsymbol{x}}_{k|k-1}$ and covariance of $\boldsymbol{P}_{k|k-1}$ can be obtained as:

$$\begin{cases} \boldsymbol{\chi}_{0,k|k-1} = \hat{\boldsymbol{x}}_{k|k-1}, \\ \boldsymbol{\chi}_{i,k|k-1} = \hat{\boldsymbol{x}}_{k|k-1} + \left(\sqrt{(n+\lambda)\boldsymbol{P}_{k|k-1}}\right)_i, i = 1,2,\cdots n, \\ \boldsymbol{\chi}_{i,k|k-1} = \hat{\boldsymbol{x}}_{k|k-1} - \left(\sqrt{(n+\lambda)\boldsymbol{P}_{k|k-1}}\right)_i, i = n+1, n+2,\cdots 2n; \end{cases} \tag{8}$$

where the weights are same as Equation (4).

The transmitting effect of the Sigma point along the nonlinear function is given as:

$$\boldsymbol{\zeta}_{i,k|k-1} = \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k|k-1}\right) + \boldsymbol{r} \quad i = 0,\cdots,2n \tag{9}$$

Then we get the posterior mean and covariance as:

$$\hat{\boldsymbol{z}}_{k|k-1} = \sum_{i=0}^{2n} \omega_i^{(m)} \cdot \boldsymbol{\zeta}_{i,k|k-1} = \sum_{i=0}^{2n} \omega_i^{(m)} \cdot \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k|k-1}\right) + \boldsymbol{r} \tag{10}$$

$$\boldsymbol{P}_{\tilde{\boldsymbol{z}}_{k|k-1}} = \sum_{i=0}^{2n} \omega_i^{(c)} \left(\boldsymbol{\zeta}_{i,k|k-1} - \hat{\boldsymbol{z}}_{k|k-1}\right) \cdot \left(\boldsymbol{\zeta}_{i,k|k-1} - \hat{\boldsymbol{z}}_{k|k-1}\right)^{\mathrm{T}} + \boldsymbol{R} \tag{11}$$

$$\boldsymbol{P}_{\tilde{\boldsymbol{x}}_{k|k-1}\tilde{\boldsymbol{z}}_{k|k-1}} = \sum_{i=0}^{2n} \omega_i^{(c)} \left(\boldsymbol{\chi}_{i,k|k-1}^{\boldsymbol{x}} - \hat{\boldsymbol{x}}_{k|k-1}\right) \cdot \left(\boldsymbol{\zeta}_{i,k|k-1} - \hat{\boldsymbol{z}}_{k|k-1}\right)^{\mathrm{T}} \tag{12}$$

STEP4. Filtering Update

$$\boldsymbol{K} = \boldsymbol{P}_{\tilde{\boldsymbol{x}}_{k|k-1}\tilde{\boldsymbol{z}}_{k|k-1}} \boldsymbol{P}_{\tilde{\boldsymbol{z}}_{k|k-1}}^{-1} \tag{13}$$

$$\hat{\boldsymbol{x}}_k = \hat{\boldsymbol{x}}_{k|k-1} + \boldsymbol{K} \cdot \left(\boldsymbol{z}_k - \hat{\boldsymbol{z}}_{k|k-1}\right) \tag{14}$$

$$\boldsymbol{P}_k = \boldsymbol{P}_{k|k-1} - \boldsymbol{K}\boldsymbol{P}_{\tilde{\boldsymbol{z}}_{k|k-1}}\boldsymbol{K}^{\mathrm{T}} \tag{15}$$

After STEP4 is finished, the algorithm returns to STEP2.

3. Noise Statistics Estimator Based on Maximum Likelihood Principle

The traditional UKF algorithm can estimate the system states successfully when the noise prior statistics is known. However, in most applications, the statistics of the noise is usually unknown. At this time, the estimation accuracy will decrease, or the filter will even diverge. Aiming at this problem, the maximum likelihood principle is employed to establish the log likelihood function with the noise statistics. Then, the EM algorithm is applied to estimate parameters online such that the log likelihood function is maximized. With the estimated noise statistics, the robustness and estimation accuracy of the UKF algorithm can be improved.

For the complex nonlinear system, the maximum likelihood principle usually cannot be used directly to acquire the parameters estimation. EM algorithm is an effective method to estimate the maximum likelihood solution in the probability model. There are two steps in EM algorithm, expectation step (E Step), and maximum step (M Step).

Assume that $\Theta = \{\boldsymbol{r}, \boldsymbol{R}, \boldsymbol{q}, \boldsymbol{Q}\}$ is the noise statistics to be estimated. The estimation of the parameters can be represented as:

$$\hat{\Theta}_{\mathrm{EM}} = \arg\max_{\Theta} \mathrm{E}\left\{\ln\left[L\left(\Theta \middle| \boldsymbol{z}_{1:k}, \boldsymbol{x}_{1:k}\right)\right]\right\} \tag{16}$$

where $\ln\left(L\left(\Theta \middle| \boldsymbol{z}_{1:k}, \boldsymbol{x}_{1:k}\right)\right)$ is the log likelihood function of the parameter Θ.

According to the definition of likelihood function, we get:

$$L\left(\Theta \middle| \boldsymbol{z}_{1:k}, \boldsymbol{x}_{1:k}\right) = p\left(\boldsymbol{z}_{1:k}, \boldsymbol{x}_{1:k} \middle| \Theta\right) \tag{17}$$

Since the system is a Markov process, we get the joint probability density function of system states and measurements as follows:

$$p\left(\boldsymbol{z}_{1:k}, \boldsymbol{x}_{1:k} \middle| \Theta\right) = p\left(\boldsymbol{x}_0 \middle| \Theta\right) \prod_{j=1}^{k} p\left(\boldsymbol{x}_j \middle| \boldsymbol{x}_{j-1}, \Theta\right) \prod_{j=1}^{k} p\left(\boldsymbol{z}_j \middle| \boldsymbol{x}_j, \Theta\right) \tag{18}$$

Assume that all the distributions are Gaussian, and then we get the following probability density function of system states initiation:

$$p\left(\boldsymbol{x}_0 \middle| \Theta\right) = (2\pi)^{-n/2} \left|\boldsymbol{P}_0\right|^{-1/2} \exp\left\{-\left[\left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right)^{\mathrm{T}} \boldsymbol{P}_0^{-1} \left(\boldsymbol{x}_0 - \hat{\boldsymbol{x}}_0\right)\right] \Big/ 2\right\} \tag{19}$$

Probability density function of the state predicting is:

$$\begin{gathered} p\left(\boldsymbol{x}_j \middle| \boldsymbol{x}_{j-1}, \Theta\right) = (2\pi)^{-n/2} \left|\boldsymbol{Q}\right|^{-1/2} \times \\ \exp\left\{-\left[\boldsymbol{x}_j - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right) - \boldsymbol{q}\right]^{\mathrm{T}} \boldsymbol{Q}^{-1} \left[\boldsymbol{x}_j - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right) - \boldsymbol{q}\right] \Big/ 2\right\} \end{gathered} \tag{20}$$

Probability density function of the measurements is:

$$\begin{gathered} p\left(\boldsymbol{z}_j \middle| \boldsymbol{x}_j, \Theta\right) = (2\pi)^{-m/2} \left|\boldsymbol{R}\right|^{-1/2} \times \\ \exp\left\{-\left[\boldsymbol{z}_j - \boldsymbol{h}\left(\boldsymbol{x}_j\right) - \boldsymbol{r}\right]^{\mathrm{T}} \boldsymbol{R}^{-1} \left[\boldsymbol{z}_j - \boldsymbol{h}\left(\boldsymbol{x}_j\right) - \boldsymbol{r}\right] \Big/ 2\right\} \end{gathered} \tag{21}$$

where operator $|\cdot|$ represents the determinant of a matrix.

From the analysis above, we can establish the log likelihood function as:

$$\ln\left(L\left(\Theta|\boldsymbol{z}_{1:k},\boldsymbol{x}_{1:k}\right)\right)=-\frac{k\left(n+m\right)+n}{2}\ln\left(2\pi\right)-\frac{1}{2}\ln\left|\boldsymbol{P}_0\right|$$
$$-\frac{1}{2}\left\{\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)^{\mathrm{T}}\boldsymbol{P}_0^{-1}\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)\right\}-\frac{k}{2}\ln\left|\boldsymbol{Q}\right|-\frac{k}{2}\ln\left|\boldsymbol{R}\right|$$
$$-\frac{1}{2}\sum_{j=1}^{k}\left\{\left[\boldsymbol{x}_j-\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)-\boldsymbol{q}\right]^{\mathrm{T}}\boldsymbol{Q}^{-1}\left[\boldsymbol{x}_j-\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)-\boldsymbol{q}\right]\right\}$$
$$-\frac{1}{2}\sum_{j=1}^{k}\left\{\left(\boldsymbol{z}_j-\boldsymbol{h}\left(\boldsymbol{x}_j\right)-\boldsymbol{r}\right)^{\mathrm{T}}\boldsymbol{R}^{-1}\left(\boldsymbol{z}_j-\boldsymbol{h}\left(\boldsymbol{x}_j\right)-\boldsymbol{r}\right)\right\} \quad (22)$$

Then, the EM algorithm can be applied for the estimation of Θ.

E Step.

By calculation the mathematical expectation of the log likelihood function of $\ln\left(L\left(\Theta|\boldsymbol{z}_{1:k},\boldsymbol{x}_{1:k}\right)\right)$, we get:

$$\mathrm{E}\left[L\left(\Theta|\boldsymbol{z}_{1:k},\boldsymbol{x}_{1:k}\right)\right]=-\frac{k\left(n+m\right)+n}{2}\ln\left(2\pi\right)-\frac{1}{2}\ln\left|\boldsymbol{P}_0\right|$$
$$-\frac{k}{2}\ln\left|\boldsymbol{Q}\right|-\frac{k}{2}\ln\left|\boldsymbol{R}\right|-\frac{1}{2}\mathrm{E}\left\{\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)^{\mathrm{T}}\boldsymbol{P}_0^{-1}\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)\right\}$$
$$-\frac{1}{2}\sum_{j=1}^{k}\mathrm{E}\left\{\left(\boldsymbol{x}_j-\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)-\boldsymbol{q}\right)^{\mathrm{T}}\boldsymbol{Q}^{-1}\left(\boldsymbol{x}_j-\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)-\boldsymbol{q}\right)\right\}$$
$$-\frac{1}{2}\sum_{j=1}^{k}\mathrm{E}\left\{\left(\boldsymbol{z}_j-\boldsymbol{h}\left(\boldsymbol{x}_j\right)-\boldsymbol{r}\right)^{\mathrm{T}}\boldsymbol{R}^{-1}\left(\boldsymbol{z}_j-\boldsymbol{h}\left(\boldsymbol{x}_j\right)-\boldsymbol{r}\right)\right\} \quad (23)$$

Since $\mathrm{E}\left[\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)^{\mathrm{T}}\boldsymbol{P}_0^{-1}\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)\right]$ only depends on the initiation of the filter, define a constant as

$$\mathrm{const}=-\frac{k\left(n+m\right)+n}{2}\ln\left(2\pi\right)-\frac{1}{2}\ln\left|\boldsymbol{P}_0\right|$$
$$-\frac{1}{2}\mathrm{E}\left[\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)^{\mathrm{T}}\boldsymbol{P}_0^{-1}\left(\boldsymbol{x}_0-\hat{\boldsymbol{x}}_0\right)\right] \quad (24)$$

By defining the mathematical expectation of the log likelihood function as objective J, since $\boldsymbol{Q}$ and $\boldsymbol{R}$ are both positively diagonal matrix, we get:

$$J=\mathrm{const}-\frac{k}{2}\ln\left|\boldsymbol{Q}\right|-\frac{k}{2}\ln\left|\boldsymbol{R}\right|$$
$$-\frac{1}{2}\sum_{j=1}^{k}\mathrm{E}\left\{tr\left[\boldsymbol{Q}^{-1}\left(\boldsymbol{x}_j-\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)-\boldsymbol{q}\right)\left(\boldsymbol{x}_j-\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)-\boldsymbol{q}\right)^{\mathrm{T}}\right]\right\}$$
$$-\frac{1}{2}\sum_{j=1}^{k}\mathrm{E}\left\{tr\left[\boldsymbol{R}^{-1}\left(\boldsymbol{z}_j-\boldsymbol{h}\left(\boldsymbol{x}_j\right)-\boldsymbol{r}\right)\left(\boldsymbol{z}_j-\boldsymbol{h}\left(\boldsymbol{x}_j\right)-\boldsymbol{r}\right)^{\mathrm{T}}\right]\right\} \quad (25)$$

M Step.

The gradient descent algorithm is applied in this step to obtain the parameter estimation that maximizing objective J:

$$\frac{\partial J}{\partial \boldsymbol{q}}=0,\frac{\partial J}{\partial \boldsymbol{Q}}=0,\frac{\partial J}{\partial \boldsymbol{r}}=0,\frac{\partial J}{\partial \boldsymbol{R}}=0 \quad (26)$$

According the above Equation, we get the estimations of $\hat{\boldsymbol{q}}_k$ and $\hat{\boldsymbol{r}}_k$ as follows:

$$\begin{cases} \hat{\boldsymbol{q}}_k = \frac{1}{k}\sum_{j=1}^{k}\left(\boldsymbol{x}_j - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)\right) \\ \hat{\boldsymbol{r}}_k = \frac{1}{k}\sum_{j=1}^{k}\left(\boldsymbol{z}_j - \boldsymbol{h}\left(\boldsymbol{x}_j\right)\right) \end{cases} \tag{27}$$

Consider that:

$$\begin{cases} \mathrm{E}\left(\boldsymbol{x}_j - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)\right) = \boldsymbol{q} \\ \mathrm{E}\left(\boldsymbol{z}_j - \boldsymbol{h}\left(\boldsymbol{x}_j\right)\right) = \boldsymbol{r} \end{cases} \tag{28}$$

Then $\hat{\boldsymbol{Q}}_k$ and $\hat{\boldsymbol{R}}_k$ can be obtained as follows:

$$\begin{cases} \hat{\boldsymbol{Q}}_k = \frac{1}{k}\sum_{j=1}^{k}\left[\left(\boldsymbol{x}_j - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right) - \boldsymbol{q}\right)\left(\boldsymbol{x}_j - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right) - \boldsymbol{q}\right)^{\mathrm{T}}\right] = \frac{1}{k}\sum_{j=1}^{k} \\ \left[\boldsymbol{x}_j\boldsymbol{x}_j^{\mathrm{T}} - \boldsymbol{q}\boldsymbol{q}^{\mathrm{T}} - \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)\boldsymbol{x}_j^{\mathrm{T}} - \boldsymbol{x}_j\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)^{\mathrm{T}} + \boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)\boldsymbol{f}\left(\boldsymbol{x}_{j-1}\right)^{\mathrm{T}}\right] \\ \hat{\boldsymbol{R}}_k = \frac{1}{k}\sum_{j=1}^{k}\left[\left(\boldsymbol{z}_j - \boldsymbol{h}\left(\boldsymbol{x}_j\right) - \boldsymbol{r}\right)\left(\boldsymbol{z}_j - \boldsymbol{h}\left(\boldsymbol{x}_j\right) - \boldsymbol{r}\right)^{\mathrm{T}}\right] = \\ \frac{1}{k}\sum_{j=1}^{k}\left[\boldsymbol{z}_j\boldsymbol{z}_j^{\mathrm{T}} - \boldsymbol{r}\boldsymbol{r}^{\mathrm{T}} - \boldsymbol{h}\left(\boldsymbol{x}_j\right)\boldsymbol{z}_j^{\mathrm{T}} - \boldsymbol{z}_j\boldsymbol{h}\left(\boldsymbol{x}_j\right)^{\mathrm{T}} + \boldsymbol{h}\left(\boldsymbol{x}_j\right)\boldsymbol{h}\left(\boldsymbol{x}_j\right)^{\mathrm{T}}\right] \end{cases} \tag{29}$$

The Sigma point of system state $\{\boldsymbol{\chi}_{i,j}\}$ and $\{\boldsymbol{\chi}_{i,j-1}\}$ are used to instead the random vector, and we can get the recursive expression as:

$$\hat{\boldsymbol{q}}_k = \frac{1}{k}\left[(k-1)\hat{\boldsymbol{q}}_{k-1} + \left(\hat{\boldsymbol{x}}_k - \sum_{i=0}^{2n}\omega_i^{(m)} \cdot \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right)\right)\right] \tag{30}$$

$$\begin{aligned} \hat{\boldsymbol{Q}}_k = \frac{1}{k}\Big\{&(k-1)\hat{\boldsymbol{Q}}_k + diag\Big[\left(\hat{\boldsymbol{x}}_k\hat{\boldsymbol{x}}_k^{\mathrm{T}} + \boldsymbol{P}_k - \hat{\boldsymbol{q}}_k\hat{\boldsymbol{q}}_k^{\mathrm{T}}\right) \\ &-\left(\sum_{i=0}^{2n}\omega_i^{(c)} \cdot \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right)\cdot\boldsymbol{\chi}_{i,k}^{\mathrm{T}}\right) - \left(\sum_{i=0}^{2n}\omega_i^{(c)} \cdot \boldsymbol{\chi}_{i,k} \cdot \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right)^{\mathrm{T}}\right) \\ &+\left(\sum_{i=0}^{2n}\omega_i^{(c)} \cdot \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right)\cdot \boldsymbol{f}\left(\boldsymbol{\chi}_{i,k-1}\right)^{\mathrm{T}}\right)\Big]\Big\} \end{aligned} \tag{31}$$

$$\hat{\boldsymbol{r}}_k = \frac{1}{k}\left[(k-1)\hat{\boldsymbol{r}}_{k-1} + \left(\boldsymbol{z}_k - \sum_{i=0}^{2n}\omega_i^{(m)} \cdot \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k}\right)\right)\right] \tag{32}$$

$$\begin{aligned} \hat{\boldsymbol{R}}_k = \frac{1}{k}\Big\{&(k-1)\hat{\boldsymbol{R}}_{k-1} + diag\Big[\left(\boldsymbol{z}_k\boldsymbol{z}_k^{\mathrm{T}} - \hat{\boldsymbol{r}}_k\hat{\boldsymbol{r}}_k^{\mathrm{T}}\right) \\ &-\left(\sum_{i=0}^{2n}\omega_i^{(m)} \cdot \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k}\right)\cdot\boldsymbol{z}_k^{\mathrm{T}}\right) - \left(\boldsymbol{z}_n \cdot \sum_{i=0}^{2n}\omega_i^{(m)} \cdot \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k}\right)^{\mathrm{T}}\right) \\ &+\left(\sum_{i=0}^{2n}\omega_i^{(c)} \cdot \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k}\right)\cdot \boldsymbol{h}\left(\boldsymbol{\chi}_{i,k}\right)^{\mathrm{T}}\right)\Big]\Big\} \end{aligned} \tag{33}$$

where the Sigma point $\{\boldsymbol{\chi}_{i,k}\}$ and $\{\boldsymbol{\chi}_{i,k-1}\}$ satisfy:

$$\begin{cases}\{\boldsymbol{\chi}_{i,k}\} \sim N(\hat{\boldsymbol{x}}_k, \boldsymbol{P}_k) \\ \{\boldsymbol{\chi}_{i,k-1}\} \sim N(\hat{\boldsymbol{x}}_{k-1}, \boldsymbol{P}_{k-1})\end{cases} \tag{34}$$

Comparing with the traditional UKF algorithm, the proposed noise statistics estimator can provide the nonlinear filter with the noise statistics information more accurate. The procedure of the adaptive UKF at k 'th moment are described as follows.

According to the noise statistics estimation of $k-1$ 'th moment $\hat{\Theta}_{k-1} = \{\hat{\boldsymbol{r}}_{k-1}, \hat{\boldsymbol{R}}_{k-1}, \hat{\boldsymbol{q}}_{k-1}, \hat{\boldsymbol{Q}}_{k-1}\}$, Equations. (3)~(15) are introduce to estimate the mean and covariance of the system states. Then, according to the estimations of the filter, Equations (30)~(33) are employed to estimate the noise statistics parameters of k 'th moment.

4. Simulations and Analysis

The following first order nonlinear system is applied for simulations:

$$\begin{cases}\boldsymbol{x}_k = 0.5\boldsymbol{x}_{k-1} + \dfrac{0.2\boldsymbol{x}_{k-1}}{1+\boldsymbol{x}_{k-1}^2} + \boldsymbol{w}_{k-1} \\ \boldsymbol{z}_k = 10\boldsymbol{x}_k + \dfrac{\boldsymbol{x}_k^2}{20} + \boldsymbol{v}_k\end{cases} \tag{48}$$

Assume that $\boldsymbol{w}$ and $\boldsymbol{v}$ are both Gaussian noise:

$$\begin{cases}\boldsymbol{w} \sim N(1.2, 0.6) \\ \boldsymbol{v} \sim N(1.0, 0.8)\end{cases} \tag{49}$$

The initial state is $\boldsymbol{x}_0 = 1$. Assume that the noise statistics of the process noise is unknown. The initial statistics of the process noise is: $\hat{\boldsymbol{q}}_0 = 0.2$, $\hat{\boldsymbol{Q}}_0 = 0.5$. The estimation effect of the proposed noise statistics estimator is shown in Figure 1 and Figure 2. It can be seen that the proposed estimator can quickly converge to the real noise statistics.

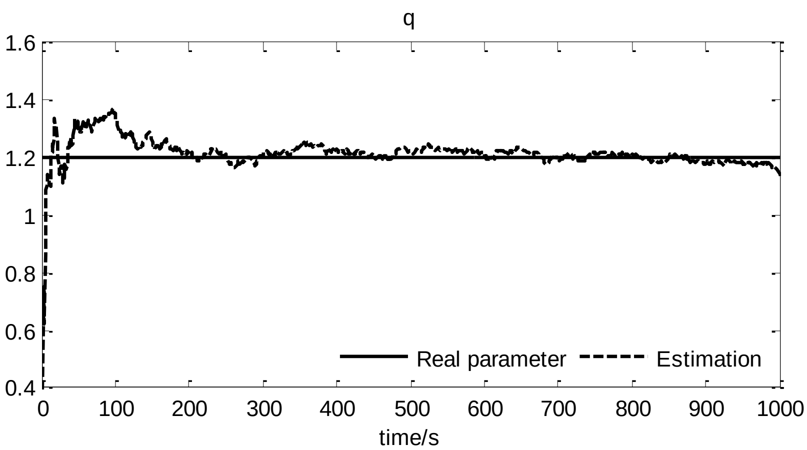

Figure 1. Mean of Process Noise Statistics

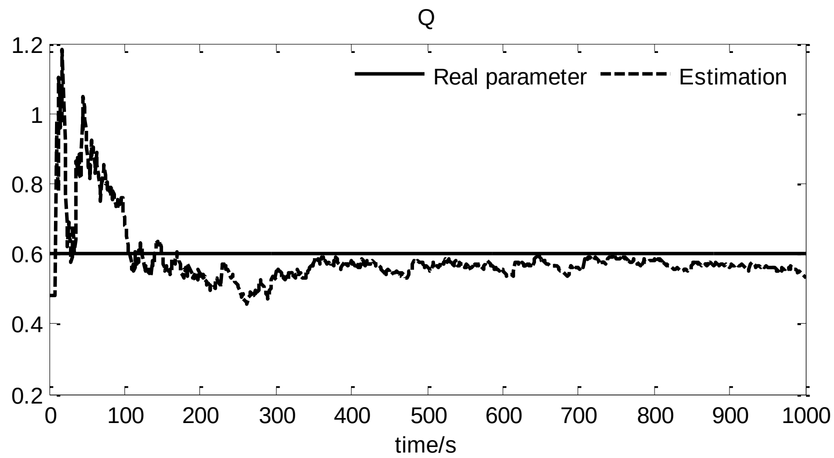

Figure 2. Covariance of Process Noise Statistics

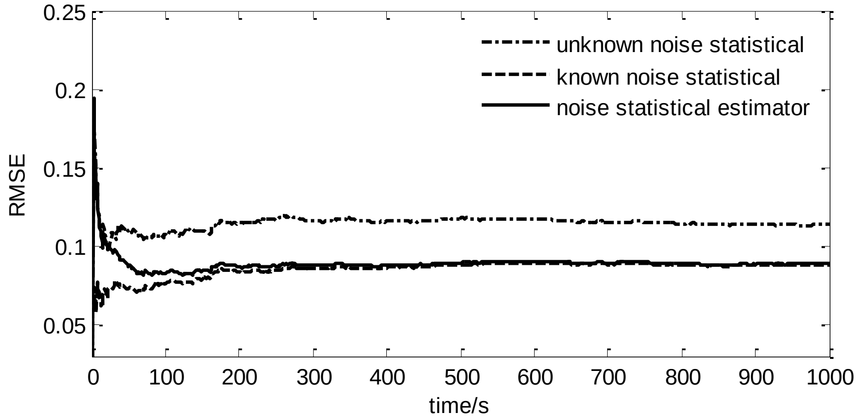

Figure 3. Root Mean Square Error

Table 1. Estimation Accuracy Contrast of Adaptive UKF and Traditional UKF Algorithm

	Mean	Covariance
UKF Algorithm(known noise statistics)	0.0001	0.0822
UKF Algorithm(unknown noise statistics)	0.1137	0.1006
Adaptive UKF	0.0006	0.0825

Both adaptive UKF proposed in this paper and traditional UKF are carried out in the simulations to show the effectiveness of the proposed algorithm. It is illustrated in Figure 3 that the proposed algorithm has higher state estimation accuracy. The state estimation accuracy of adaptive UKF is much higher than that of the traditional UKF algorithm with unknown noise statistics. Notice that the proposed estimator can converge at about 100 filter periods. Consequently, the accuracy of the proposed adaptive UKF with unknown noise statistics is similar to the traditional UKF with known information after the noise statistics estimator converges. The accuracy comparison is shown in Table 1.

5. Conclusions

In this paper, the nonlinear filtering problem is investigated with unknown noise statistics. An adaptive UKF algorithm with noise statistics estimator is proposed. The proposed strategy can estimate the noise statistics as well as the system states. Simulations show that comparing with the traditional UKF algorithm, the proposed strategy can estimate the system states more accurately.

References

[1] M. Norgaard, N. K. Poulsen and O. Ravn, "New developments in state estimation for nonlinear systems", Automatica, vol. 36, no. 11, **(2000)**.

[2] E. Belviken, P. J. Acklam and M. Carlo, "filters for nonlinear state estimation", Automatica, vol. 37, no. 1, (**2001**).

[3] S. Julier, J. K. Uhlmann and H. F. D. Whyte, "A new approach for filtering nonlinear system", Proceeding of the American Control Conference, Seattle, WA,USA, **(1995)**.

[4] K. Xiong, H. Y. Zhang and C. W. Chan, "Performance evaluation of UKF based nonlinear filtering", Automatica, vol. 42, no. 2, **(2006)**.

[5] L. A. Lesmes, S. T. Jeon and Z. L. Lu, "Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method", Vision Research, vol. 46, no. 19, **(2006)**.

[6] S. Parthasarathy and C. Balaji, "Estimation of parameters in multi-mode heat transfer problems using Bayesian inference – Effect of noise and a priori", International Journal of Heat and Mass Transfer, vol. 51, no. 9-10, **(2008)**.

[7] A. P. Age and G. W. Husa, "Adaptive Filtering with Unknown Prior Statistics", Proceedings of Joint Automatic Control Conference, **(1969)**.

[8] E. Saatci and A. Akan, "Respiratory parameter estimation in non-invasive ventilation based on generalized Gaussian noise models", Signal Processing, vol. 90, no. 2, **(2010)**.

[9] B. M. Akesson, J. B. Jorgensen and N. K. Poulsen, "A generalized autocovariance least-squares method fror Kalman filter tuning", Journal of Process Control, vol. 18, no. 7-8, **(2008)**.

[10] B. J. Odelson, M. R. Rajamani and J. B. Rawlings, "A new autocovariance least-squares method for estimating noise covariances", Automatica, vol. 42, no. 2, **(2006)**.

[11] R. Mehra, "Approaches to Adaptive Filtering", IEEE Transactions on Automatic Control, vol. 17, no. 5, **(1972)**.

[12] K. Myers and B. Tapley, "Adaptive Sequential Estimation with Unknown Noise Statistics", IEEE Transactions on Automatic Control, vol. 21, no. 4, **(1976)**.

Author

Li Guo, he received the B.S. degree in Mechanical Manufacturing and Automation from Dalian University of Technology, China, in 2002, the M.S. degree in Mechanical Manufacturing and Automation from Dalian University of Technology, China, in 2005. Currently, he is an associate professor in the School of Information Science and Engineering, Dalian Polytechnic University, China. His research interests include information retrieval, optimization algorithm, data mining and intelligent computing.

