

交互式容积卡尔曼滤波及其应用

于 滢, 宋申民, 王 硕

(哈尔滨工业大学 航天学院, 哈尔滨 150001)

摘要: 针对非线性系统模型参数未知情况下的状态估计问题, 提出一种融合极大后验估计的交互式容积卡尔曼滤波算法(InCKF). 该算法利用二阶斯特林插值公式和无迹变换对非线性函数的近似思想, 实现对模型未知参数的确定, 从而使滤波算法摆脱对模型参数精确已知的依赖, 并通过容积卡尔曼滤波算法完成状态估计和量测更新. 仿真结果表明, 相比于经典的参数扩维方法, InCKF算法具有更高的精度和更强的数值稳定性.

关键词: 非线性滤波; 容积卡尔曼滤波; 极大后验估计; 模型不确定性

中图分类号: V448.2

文献标志码: A

Interaction cubature Kalman filter and its application

YU Han, SONG Shen-min, WANG Shuo

(School of Astronautics, Harbin Institute of Technology, Harbin 150001, China. Correspondent: YU Han, E-mail: yuhanihit@163.com)

Abstract: For the problem of states estimation of nonlinear systems with respect to unknown parameters, a filter named interaction cubature Kalman filter(InCKF) is proposed. The novel filter is consisted of multiple concurrent CKFs interlacing with a maximum posteriori(MAP) estimator. By taking advantage of special properties of second order of Stirling's interpolation and unscented transformation to approximate nonlinear functions, the unknown parameters are estimated and the performance of InCKF does not depend on the precision of model parameters. Furthermore, the states of system are estimated by using the cubature Kalman filter(CKF). The simulation results show that the InCKF is more accurate and stabilized than the classical method of state augmentation in the situation that model parameters are unknown.

Keywords: nonlinear filter; cubature Kalman filter; maximum a posteriori; model uncertainty

0 引 言

卡尔曼滤波可以得到最小方差意义下的最优解, 但仅适用于线性系统. 然而, 社会生产生活中所涉及的问题往往是非线性的. 例如, 空间卫星交会对接^[1]、无人机导航^[2]和计量经济分析^[3]等问题. 因此, 对非线性系统状态估计问题的研究具有重要的理论意义和实际应用价值.

扩展卡尔曼滤波(EKF)^[4]由于其简洁的形式和高效的计算性能, 已成为在实际系统中应用最为广泛的一种非线性滤波. 然而, 由于一阶泰勒展开对非线性函数的近似精度偏低, 在处理强非线性函数时, EKF滤波精度较差, 且数值稳定性较差. 此外, 在估计高维系统的状态时, 线性化过程中所需的雅克比矩阵的计算十分困难, 且某些系统的雅克比矩阵并不存在^[5]. 针对EKF无法很好地解决强非线性系统的

状态估计问题, 研究人员相继提出了一系列非线性滤波算法, 其中最具代表性的算法是无迹卡尔曼滤波算法(UKF)和粒子滤波算法(PF). Julier等^[5]从“对非线性函数的概率密度分布的近似比对非线性函数本身的近似更容易”的角度出发, 提出了利用无迹变换(UT)后的采样点逼近非线性函数概率密度分布的UKF滤波算法, 采用的UT变换能够匹配任意非线性函数的三阶泰勒展开, 即达到对非线性函数的三阶逼近精度. 张勇刚等^[6]在五阶容积变换的基础上, 给出了高阶无迹变换的解析解, 并提出了一种高阶无迹卡尔曼滤波算法. 与UKF等确定采样型滤波器不同的是, PF基于蒙特卡罗仿真对先验分布进行随机采样^[7]. 由于不受限于噪声统计特性为高斯白噪声的假设, 对非高斯动态系统的状态估计问题具有较高的精度^[8], 且只要粒子数目足够多, 便可达到对非线性系

收稿日期: 2014-06-09; 修回日期: 2014-12-31.

基金项目: 国家自然科学基金项目(61174037); 国家自然科学基金创新群体项目(61021002).

作者简介: 于滢(1985—), 男, 博士生, 从事航天器姿态确定、非线性滤波的研究; 宋申民(1968—), 男, 教授, 博士生导师, 从事鲁棒控制、智能优化与智能控制等研究.

统的任意阶估计精度. 然而, PF 滤波算法存在的样本贫化问题和算法实时性等问题, 限制了滤波算法的应用, 难以满足实际工程的需要. 在贝叶斯滤波框架下, 对状态后验概率分布的估计本质上是对高维积分的计算, 然而直接计算非线性函数的高维积分是十分困难的. Arasaratnam 等^[9]提出以三阶容积法则来近似加权高斯积分的容积卡尔曼滤波(CKF), 该算法采用等权值(系统状态维数)的容积点进行非线性变换, 逼近后验概率密度分布, 解决了 UKF 在处理高维系统非线性估计时, 数值稳定性较差甚至发散的问题.

当系统存在未建模动态、模型参数未知或噪声统计特性不准确时, CKF 和其他贝叶斯估计算法无法解决上述问题. Moradkhani 等^[10]针对水文模型参数不确定性的问题, 提出了双状态-参数估计的集合卡尔曼滤波器, 使系统状态和模型参数能够同时估计. Aghili 等^[11]将空间自由漂浮卫星的转动惯量这一未知参数与位置、姿态等信息一并作为系统状态估计.

系统状态维数的增加会导致滤波器数值稳定性下降, 甚至造成滤波发散的严重后果. 针对模型参数未知的问题, 本文提出一种融合极大后验估计(MAP)的交互式容积卡尔曼滤波算法(InCKF), 采用 MAP 对模型未知参数进行动态识别, 从而解决模型参数未知情况下, 滤波算法性能下降的问题. 同时提出基于二阶斯特林插值公式(SI2)和 UT 变换两种形式的条件概率密度估计算法, 以实现关于模型未知参数的 MAP 估计. 仿真结果表明, 相比于经典的参数扩维方法, InCKF 算法具有更高的精度和更强的数值稳定性.

1 CKF

考虑具有加性噪声的非线性离散系统

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}) + \mathbf{v}_{k-1}, \quad (1)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{n}_k. \quad (2)$$

其中: $\mathbf{x}_k \in \mathbf{R}^n$ 为系统在 k 时刻的状态; $\mathbf{y}_k \in \mathbf{R}^m$ 为系统在 k 时刻的量测值; $\mathbf{u}_k \in \mathbf{R}^u$ 为系统输入; \mathbf{v}_{k-1} 和 \mathbf{n}_k 为不相关零均值高斯白噪声, 其协方差矩阵分别为 \mathbf{Q}_{k-1} 和 \mathbf{R}_{k-1} . CKF 滤波算法由时间预测和量测更新两部分组成, 具体流程如下^[9]:

1) 时间预测为

$$\hat{\mathbf{x}}_k^- = \frac{1}{2n} \sum_{i=1}^{2n} \mathbf{X}_{i,k}^*, \quad (3)$$

$$\mathbf{P}_k^- = \frac{1}{2n} \sum_{i=1}^{2n} \mathbf{X}_{i,k}^* \mathbf{X}_{i,k}^{*\top} - \hat{\mathbf{x}}_k^- \hat{\mathbf{x}}_k^{-\top} + \mathbf{Q}_{k-1}. \quad (4)$$

其中

$$\mathbf{X}_{i,k}^* = \mathbf{f}(\mathbf{X}_{i,k-1}, \mathbf{u}_{k-1}), \quad i = 1, 2, \dots, 2n; \quad (5)$$

$$\mathbf{X}_{i,k-1} = \mathbf{S}_{k-1} \boldsymbol{\xi}_i + \hat{\mathbf{x}}_{k-1}, \quad i = 1, 2, \dots, 2n; \quad (6)$$

$$\mathbf{P}_{k-1} = \mathbf{S}_{k-1} \mathbf{S}_{k-1}^\top, \quad (7)$$

\mathbf{P}_{k-1} 为 $k-1$ 时刻的协方差阵, $\boldsymbol{\xi}_i$ 为第 i 个容积点.

2) 量测更新为

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_k^-), \quad (8)$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{\hat{\mathbf{z}}\hat{\mathbf{z}},k} \mathbf{K}_k^\top. \quad (9)$$

其中

$$\mathbf{K}_k = \mathbf{P}_{\mathbf{x}\mathbf{z},k} \mathbf{P}_{\hat{\mathbf{z}}\hat{\mathbf{z}},k}^{-1}; \quad (10)$$

$$\mathbf{P}_{\hat{\mathbf{z}}\hat{\mathbf{z}},k} = \frac{1}{2n} \sum_{i=1}^{2n} \mathbf{Z}_{i,k} \mathbf{Z}_{i,k}^\top - \hat{\mathbf{z}}_k^- \hat{\mathbf{z}}_k^{-\top} + \mathbf{R}_k; \quad (11)$$

$$\mathbf{P}_{\mathbf{x}\mathbf{z},k} = \frac{1}{2n} \sum_{i=1}^{2n} \mathbf{X}_{i,k} \mathbf{Z}_{i,k}^\top - \hat{\mathbf{x}}_k^- \hat{\mathbf{z}}_k^{-\top}; \quad (12)$$

$$\hat{\mathbf{z}}_k^- = \frac{1}{2n} \sum_{i=1}^{2n} \mathbf{Z}_{i,k}; \quad (13)$$

$$\mathbf{Z}_{i,k} = \mathbf{h}(\mathbf{X}_{i,k}, \mathbf{u}_k), \quad i = 1, 2, \dots, 2n; \quad (14)$$

$$\mathbf{X}_{i,k} = \mathbf{S}_k^- \boldsymbol{\xi}_i + \hat{\mathbf{x}}_k^-, \quad i = 1, 2, \dots, 2n; \quad (15)$$

$$\mathbf{P}_k^- = \mathbf{S}_k^- \mathbf{S}_k^{-\top}. \quad (16)$$

2 InCKF

引入 InCKF 算法以解决系统模型中含有未知参数的状态估计问题. 考虑含有未知参数的非线性离散系统

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{u}_{k-1}, \boldsymbol{\theta}) + \mathbf{v}_{k-1}, \quad (17)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, \mathbf{u}_k, \boldsymbol{\theta}) + \mathbf{n}_k. \quad (18)$$

其中: $\boldsymbol{\theta} \in \mathbf{R}^{n_\theta}$ 为离散系统的未知参数, 其他参数与式(1)和(2)定义相同. InCKF 算法由多个 CKF 滤波器和一个极大后验估计器构成. 其中: CKF 滤波器的状态估计结果提供给极大后验估计器, 以确定当前时刻未知参数 $\boldsymbol{\theta}$ 的估计值; 极大后验估计器得到关于参数 $\boldsymbol{\theta}$ 的估计值后, 再传递给 CKF 滤波器, 以便完成下一时刻的状态估计. 在所设计的 InCKF 算法中, 采用多个 CKF 滤波器并行工作的方式, 每个滤波器获得系统状态的条件期望近似值, 即

$$\hat{\mathbf{x}}_k(\boldsymbol{\theta}_j) \approx \mathbf{E}\{\mathbf{x}_k | \mathbf{y}_{1:k}, \boldsymbol{\theta}_j\}. \quad (19)$$

其中: $\mathbf{y}_{1:k} = [\mathbf{y}_1^\top \mathbf{y}_2^\top \dots \mathbf{y}_k^\top]^\top$ 为 k 时刻的观测值, $\boldsymbol{\theta}_j$ 为当前时刻下对参数 $\boldsymbol{\theta}$ 的第 j 个假设. 采用极大后验估计器, 获得对当前时刻参数 $\boldsymbol{\theta}$ 的估计. 由贝叶斯估计可得到关于参数 $\boldsymbol{\theta}$ 的后验概率密度估计为

$$p(\boldsymbol{\theta}_j | \mathbf{y}_{1:k}) \propto p(\mathbf{y}_{1:k} | \boldsymbol{\theta}_j) p(\boldsymbol{\theta} = \boldsymbol{\theta}_j). \quad (20)$$

其中: $p(\boldsymbol{\theta})$ 为未知参数 $\boldsymbol{\theta}$ 的先验概率密度; $p(\mathbf{y}_{1:k} | \boldsymbol{\theta}_j)$ 为观测值关于假设 $\boldsymbol{\theta}_j$ 的条件概率密度, 其根据乘法公式分解为

$$p(\mathbf{y}_{1:k} | \boldsymbol{\theta}_j) = p(\mathbf{y}_1 | \boldsymbol{\theta}_j) \prod_{i=2}^k p(\mathbf{y}_i | \boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}). \quad (21)$$

式(21)等号右边的因子 $p(\mathbf{y}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1})$ 展开为

$$\begin{aligned} p(\mathbf{y}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}) &= \\ \int \int p(\mathbf{y}_i|\mathbf{x}_i)p(\mathbf{x}_i|\mathbf{x}_{i-1}, \boldsymbol{\theta}_j)p(\mathbf{x}_{i-1}|\mathbf{y}_{1:i-1})d\mathbf{x}_i d\mathbf{x}_{i-1} &= \\ \int p(\mathbf{y}_i|\mathbf{x}_i)p(\mathbf{x}_i|\mathbf{y}_{1:i-1}, \boldsymbol{\theta}_j)d\mathbf{x}_i &= \\ E\{p(\mathbf{y}_i|\mathbf{x}_i)|\mathbf{y}_{1:i-1}, \boldsymbol{\theta}_j\}. \end{aligned} \quad (22)$$

由上述分析可知, 只要获得了 $p(\mathbf{y}_{1:k}|\boldsymbol{\theta}_j)$ 的估计值, 便可以采用极大后验估计的方式, 得到关于未知参数 $\boldsymbol{\theta}$ 的估计值. 本文提出两种方法估计条件概率密度 $p(\mathbf{y}_{1:k}|\boldsymbol{\theta}_j)$.

2.1 基于二阶斯特林插值的概率密度估计

二阶斯特林插值公式^[12]如下: 考虑非线性函数 $\mathbf{y} = \mathbf{f}(\mathbf{x})$, 根据 SI2, 在点 $\hat{\mathbf{x}}$ 处, 有

$$\begin{aligned} \mathbf{y} &\approx \\ \mathbf{f}(\hat{\mathbf{x}}) + \sum_{d=1}^n \Delta x_d \frac{\mathbf{f}(\hat{\mathbf{x}} + h\mathbf{S}_d) - \mathbf{f}(\hat{\mathbf{x}} - h\mathbf{S}_d)}{2h} + \\ \sum_{d=1}^n (\Delta x_d)^2 \frac{\mathbf{f}(\hat{\mathbf{x}} + h\mathbf{S}_d) + \mathbf{f}(\hat{\mathbf{x}} - h\mathbf{S}_d) - 2\mathbf{f}(\hat{\mathbf{x}})}{2h^2}. \end{aligned} \quad (23)$$

其中: Δx_d 为向量 $\Delta \mathbf{x} = \mathbf{x} - \hat{\mathbf{x}}$ 的第 d 个元素; \mathbf{S}_d 为协方差矩阵平方根分解后的第 d 列向量; h 等于插值区间的一半. 根据 SI2, 式(22)的概率密度 $p(\mathbf{y}_i|\mathbf{x}_i)$ 可以近似为

$$\begin{aligned} \hat{p}(\mathbf{y}_i|\mathbf{x}_i) &= \\ p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)) + \sum_{d=1}^n (\Delta x_i)_d \left(\frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) + h\mathbf{S}_d)}{2h} - \right. \\ \left. \frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) - h\mathbf{S}_d)}{2h} \right) + \\ \sum_{d=1}^n (\Delta x_i)_d^2 \left(\frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) + h\mathbf{S}_d)}{2h^2} + \right. \\ \left. \frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) - h\mathbf{S}_d) - 2p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j))}{2h^2} \right). \end{aligned} \quad (24)$$

其中: $\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)$ 为在当前时刻 i 未知参数等于 $\boldsymbol{\theta}_j$ 条件下的状态预测值, $\Delta \mathbf{x}_i = \mathbf{x}_i - \hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)$. 假设 $\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) \approx E\{\mathbf{x}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}\}$, 将式(24)代入(22), 得到

$$\begin{aligned} p(\mathbf{y}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}) &\approx E\{\hat{p}(\mathbf{y}_i|\mathbf{x}_i)|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}\} = \\ p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)) + E\left\{ \sum_{d=1}^n (\Delta x_i)_d \left(\frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) + h\mathbf{S}_d)}{2h} - \right. \right. \\ \left. \left. \frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) - h\mathbf{S}_d)}{2h} \right) \middle| \boldsymbol{\theta}_j, \mathbf{y}_{1:i-1} \right\} + \\ E\left\{ \sum_{d=1}^n (\Delta x_i)_d^2 \left(\frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) + h\mathbf{S}_d)}{2h^2} + \right. \right. \\ \left. \left. \frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j) - h\mathbf{S}_d)}{2h^2} \right) \middle| \boldsymbol{\theta}_j, \mathbf{y}_{1:i-1} \right\} + \\ E\left\{ \text{Tr} \left[\frac{-p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j))}{h^2} (\Delta \mathbf{x}_i)(\Delta \mathbf{x}_i)^T \right] \middle| \boldsymbol{\theta}_j, \mathbf{y}_{1:i-1} \right\} = \end{aligned}$$

$$p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)) - \frac{p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j))}{h^2} \text{Tr}[\text{cov}(\mathbf{x}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1})]. \quad (25)$$

如果 $\text{Tr}[\text{cov}(\mathbf{x}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1})]$ 足够小, 则有

$$p(\mathbf{y}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}) \approx p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)). \quad (26)$$

因此, 式(21)近似为

$$p(\mathbf{y}_{1:k}|\boldsymbol{\theta}_j) \approx \prod_{i=1}^k p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)). \quad (27)$$

令

$$\hat{p}_{\text{SI2}}(\mathbf{y}_{1:k}|\boldsymbol{\theta}_j) := \prod_{i=1}^k p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)), \quad (28)$$

对式(28)两边取对数, 并代入式(18), 得到

$$\begin{aligned} \log \hat{p}_{\text{SI2}}(\mathbf{y}_{1:k}|\boldsymbol{\theta}_j) &= \sum_{i=1}^k \log p(\mathbf{y}_i|\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)) = \\ \sum_{i=1}^k \log p(\mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j))) &= \\ -\frac{k}{2} \log |2\pi \mathbf{R}| - \\ \frac{1}{2} \sum_{i=1}^k (\mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j)))^T \mathbf{R}^{-1} (\mathbf{y}_i - \mathbf{h}(\hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j))). \end{aligned} \quad (29)$$

2.2 基于无迹变换的概率密度估计

无迹变换^[5]如下: 考虑非线性函数 $\mathbf{y} = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in \mathbf{R}^n$. 设存在如下集合 $\{\mathcal{X}_d\}$ 满足

$$\begin{aligned} \mathcal{X}_0 &= \hat{\mathbf{x}}; \\ \mathcal{X}_d &= \hat{\mathbf{x}} + (\sqrt{(n+\lambda)\mathbf{P}})_d, \quad d = 1, 2, \dots, n; \\ \mathcal{X}_d &= \hat{\mathbf{x}} - (\sqrt{(n+\lambda)\mathbf{P}})_{d-n}, \\ & \quad d = n+1, n+2, \dots, 2n. \end{aligned} \quad (30)$$

其中: $\hat{\mathbf{x}}$ 为 \mathbf{x} 的均值; $(\sqrt{(n+\lambda)\mathbf{P}})_d$ 为矩阵第 d 列向量; \mathbf{P} 为协方差阵; 尺度算子 λ 取 $\lambda = \alpha^2(n+k) - n$, α 一般取较小的整数 (如 $1 \leq \alpha \leq 10^{-4}$), k 一般取 $3-n$. 因此, 对非线性函数 $\mathbf{y} = \mathbf{f}(\mathbf{x})$ 的近似值为

$$\hat{\mathbf{y}} \approx \sum_{d=0}^{2n} W_d \mathbf{f}(\mathcal{X}_d). \quad (31)$$

其中 $W_0 = \frac{\lambda}{n+\lambda}$, $W_d = \frac{1}{2(n+\lambda)}$.

根据 UT 变换, 式(21)中的因子 $p(\mathbf{y}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1})$ 可近似为

$$\begin{aligned} p(\mathbf{y}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1}) &= \\ \int \int p(\mathbf{y}_i|\mathbf{x}_i)p(\mathbf{x}_i|\mathbf{x}_{i-1}, \boldsymbol{\theta}_j)p(\mathbf{x}_{i-1}|\mathbf{y}_{1:i-1})d\mathbf{x}_i d\mathbf{x}_{i-1} &= \\ \int p(\mathbf{y}_i|\mathbf{x}_i)p(\mathbf{x}_i|\boldsymbol{\theta}_j, \mathbf{y}_{1:i-1})d\mathbf{x}_i &= \\ \int p(\mathbf{y}_i|\mathbf{x}_i)\mathcal{N}(\mathbf{x}_i; \hat{\mathbf{x}}_i^-(\boldsymbol{\theta}_j), \mathbf{P}_i^-)d\mathbf{x}_i \approx \\ \sum_{d=0}^{2n} W_d p(\mathbf{y}_i|\mathcal{X}_d). \end{aligned} \quad (32)$$

其中

$$\begin{aligned} \mathcal{X}_0 &= \hat{\mathbf{x}}_i^-(\theta_j); \\ \mathcal{X}_d &= \hat{\mathbf{x}}_i^-(\theta_j) + (\sqrt{(n+\lambda)\mathbf{P}_k^-})_d, \quad d = 1, 2, \dots, n; \\ \mathcal{X}_d &= \hat{\mathbf{x}}_i^-(\theta_j) - (\sqrt{(n+\lambda)\mathbf{P}_k^-})_{d-n}, \\ & \quad d = n+1, n+2, \dots, 2n; \end{aligned} \quad (33)$$

$\mathcal{N}(\cdot, \cdot)$ 为高斯概率密度. 式(21)近似为

$$\begin{aligned} p(\mathbf{y}_{1:k}|\theta_j) &\approx \prod_{i=1}^k \left(\sum_{d=0}^{2n} W_d p(\mathbf{y}_i|\mathcal{X}_d) \right) = \\ &\prod_{i=1}^k \left(\sum_{d=0}^{2n} W_d p(\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d)) \right) = \\ &\prod_{i=1}^k \left(\sum_{d=0}^{2n} W_d \frac{1}{\sqrt{|2\pi\mathbf{R}|}} e^{-\frac{(\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d))^T \mathbf{R}^{-1} (\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d))}{2}} \right) = \\ &\prod_{i=1}^k \frac{1}{\sqrt{|2\pi\mathbf{R}|}} \left(\sum_{d=0}^{2n} W_d e^{-\frac{(\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d))^T \mathbf{R}^{-1} (\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d))}{2}} \right). \end{aligned} \quad (34)$$

令 $M_{\text{UT}} := -\frac{(\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d))^T \mathbf{R}^{-1} (\mathbf{y}_i - \mathbf{h}(\mathcal{X}_d))}{2}$, 对 $e^{(\cdot)}$ 取二阶泰勒近似, 有

$$\begin{aligned} p(\mathbf{y}_{1:k}|\theta_j) &\approx \\ &\prod_{i=1}^k \frac{1}{\sqrt{|2\pi\mathbf{R}|}} \left(\sum_{d=0}^{2n} W_d \left(1 + M_{\text{UT}} + \frac{M_{\text{UT}}^2}{2} \right) \right). \end{aligned} \quad (35)$$

令

$$\begin{aligned} \hat{p}_{\text{UT}}(\mathbf{y}_{1:k}|\theta_j) &:= \\ &\prod_{i=1}^k \frac{1}{\sqrt{|2\pi\mathbf{R}|}} \left(\sum_{d=0}^{2n} W_d \left(1 + M_{\text{UT}} + \frac{M_{\text{UT}}^2}{2} \right) \right), \end{aligned} \quad (36)$$

对式(36)两边取对数, 得到

$$\begin{aligned} \log \hat{p}_{\text{UT}}(\mathbf{y}_{1:k}|\theta_j) &= \\ \log \left(\prod_{i=1}^k \frac{1}{\sqrt{|2\pi\mathbf{R}|}} \left(\sum_{d=0}^{2n} W_d \left(1 + M_{\text{UT}} + \frac{M_{\text{UT}}^2}{2} \right) \right) \right) &= \\ \sum_{i=1}^k \frac{1}{\sqrt{|2\pi\mathbf{R}|}} \left(\sum_{d=0}^{2n} W_d \left(1 + M_{\text{UT}} + \frac{M_{\text{UT}}^2}{2} \right) \right). \end{aligned} \quad (37)$$

2.3 InCKF

采用上述提出的概率密度估计方法, 得到未知参数 θ 的极大后验概率密度估计为

$$\begin{aligned} \hat{\theta}_{\text{MAP}} &= \\ \arg \max_{\theta_j \in \theta} \log \hat{p}_{(\cdot)}(\mathbf{y}_{1:k}|\theta_j) + \log p(\theta = \theta_j), \end{aligned} \quad (38)$$

其中 $\log \hat{p}_{(\cdot)}(\mathbf{y}_{1:k}|\theta_j)$ 表示式(29)和(37). 此外, 如果参数 θ 的先验概率密度 $p(\theta)$ 精度较低, 则极大后验概率密度估计修正为

$$\begin{aligned} \hat{\theta}_{\text{MAP}} &= \\ \arg \max_{\theta_j \in \theta} \log \hat{p}_{(\cdot)}(\mathbf{y}_{1:k}|\theta_j) + \log \hat{p}_{(\cdot)}(\theta_{\text{MAP}}|\mathbf{y}_{1:k}). \end{aligned} \quad (39)$$

其中

$$\hat{p}_{(\cdot)}(\theta_{\text{MAP}}|\mathbf{y}_{1:k}) = \hat{p}_{(\cdot)}(\mathbf{y}_{1:k}|\theta_{\text{MAP}})p(\theta = \theta_{\text{MAP}}), \quad (40)$$

$\hat{p}_{(\cdot)}(\mathbf{y}_{1:k}|\theta_{\text{MAP}})$ 为式(28)和(36). 综上所述, InCKF 算法流程如下.

Step 1: 参数初始化. 取状态估计值 $\hat{\mathbf{x}}_0 = \mathbf{0}$, $\hat{\theta}_{\text{MAP}} = I$, I 为单位阵, 算法运行次数设定为 M , 计数值 k 记为 1.

Step 2: 从未知参数 θ 的先验概率密度 $p(\theta)$ 中采样 N_θ 次, 获取关于参数 θ 的 N_θ 个假设值 θ_j .

Step 3: 令 $\theta_0 := \hat{\theta}_{\text{MAP}}$, 计数值 j 记为 $j = 0$.

Step 4: 将假设值 θ_j 代入 CKF 滤波器中, 获取 k 时刻的状态预测值 $\hat{\mathbf{x}}_k^-(\theta_j)$ 和估计值 $\hat{\mathbf{x}}_k(\theta_j)$, 根据式(29)(或(37))计算 $\log \hat{p}(\mathbf{y}_{1:k}|\theta_j)$, $j = j + 1$.

Step 5: 判断 $j < N_\theta + 1$ 是否成立, 若成立, 则返回 Step 4, 否则执行 Step 6.

Step 6: 根据式(39)计算参数 θ 的估计值, 根据式(40)修正先验概率密度 $p(\theta)$, 即

$$p(\theta) = \hat{p}_{(\cdot)}(\theta_{\text{MAP}}|\mathbf{y}_{1:k}),$$

取滤波算法输出为 $\hat{\mathbf{x}}_k(\hat{\theta}_{\text{MAP}})$, $k = k + 1$.

Step 7: 判断 $k < M$ 是否成立, 若成立, 则转至 Step 2, 否则算法结束.

根据上述步骤, 整理算法流程如图 1 所示.

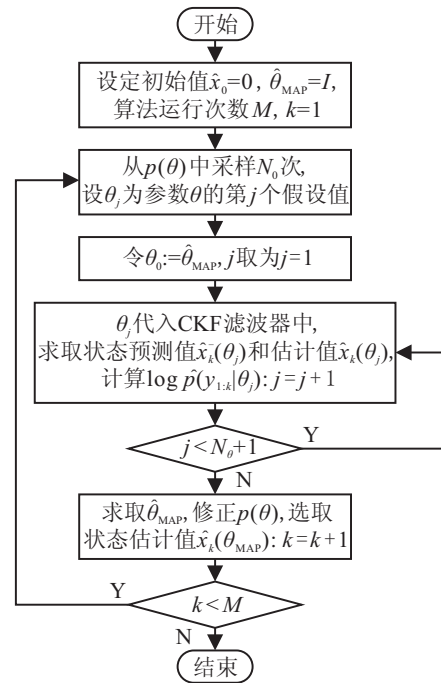


图 1 InCKF 算法流程

3 仿真分析

为了验证本文提出的 InCKF 算法的有效性和可行性, 选用在非线滤波领域中的典型测试模型——单变量非静态增长模型 (UNGM)^[13-15] 进行仿真实验. 仿真中, 对模型参数精确已知时采用的 CKF 算法、模型参数未知时采用 EKF 算法框架下的参数扩维方法和 InCKF 算法下的滤波结果进行比较. 对模型未知参

数的辨识分定常和时变两种情况进行仿真分析。

情况1 模型未知参数定常情况下仿真。

仿真中采用的UNGM动态过程为

$$x_k = ax_{k-1} + b \frac{x_{k-1}}{1 + x_{k-1}^2} + c \cos(1.2(k-1)) + w_{k-1}, \quad (41)$$

$$y_k = dx_k^2 + v_k. \quad (42)$$

其中: w_{k-1} 和 v_k 为零均值高斯白噪声, 其方差分别为 0.1 和 1; 未知参数 $\theta = [a \ b \ c \ d]^T$, 其真实值 $\theta^* = [a^* \ b^* \ c^* \ d^*] = [0.5 \ 25 \ 8 \ 0.05]^T$, 先验概率密度 $p(\theta)$ 满足均值为 $[0.4 \ 24 \ 9 \ 0.07]^T$, 方差为 $\text{diag}(0.1, 10, 1, 0.001)$ 的高斯分布. 定义模型参数估计误差为

$$\Delta\theta = \sqrt{(a^* - \hat{a})^2 + (b^* - \hat{b})^2 + (c^* - \hat{c})^2 + (d^* - \hat{d})^2}. \quad (43)$$

初始状态真值 $x = 0.1$, 滤波器初始值 $\hat{x} = 0.1$, 协方差阵初始值 $P_0 = I_{1 \times 1}$, InCKF 算法参数 $N_\theta = 5$.

图2~图5分别为参数 θ 定常情况下, 状态真实值和滤波估计结果比较、状态估计误差、状态估计误差范数、模型参数估计误差. 由图2和图3可见, 在 $t = 30$ 和 $t = 82$ 附近, 系统非线性特性较为严重, 4种滤波算法对状态的估计均出现较大误差. 对比估计结果可知, 扩维EKF算法性能较差, 在整个滤波周期中

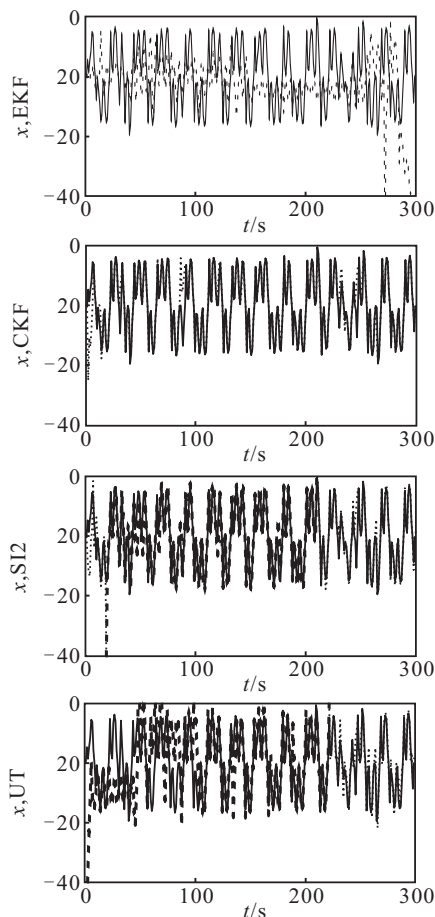


图2 参数 θ 定常情况下的状态真实值和滤波估计结果比较

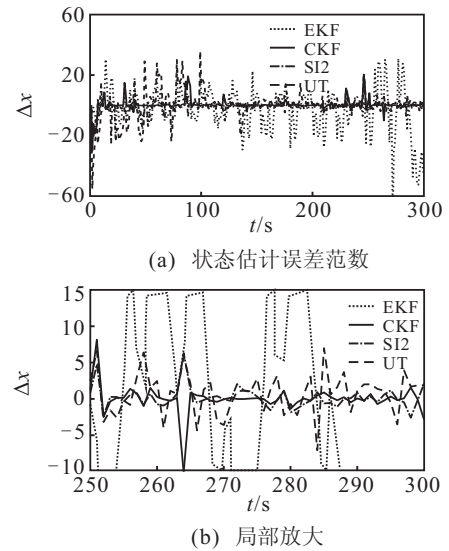


图3 参数 θ 定常情况下的状态估计误差及局部放大

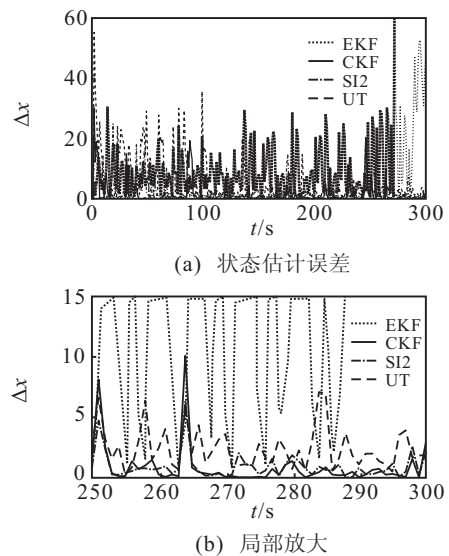


图4 参数 θ 定常情况下的状态估计误差范数及局部放大

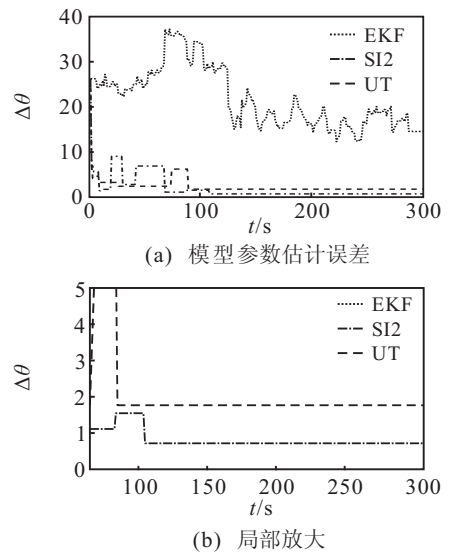


图5 参数 θ 定常情况下的模型参数估计误差及局部放大

均出现较大的估计误差, 尤其在 $t = 90 \sim 100$ 的区间内, 性能最差, 因此, 扩维EKF算法不适合解决强非线性

性模型中存在参数未知情况下的状态估计问题。

由图2、图3和图5可见, 在状态估计前段, 即 $t=30$ 之前, 由于 InCKF 算法对模型参数的估计还未处于稳定, 基于 SI2 的 InCKF(SI2-InCKF) 和基于 UT 变换的 InCKF(UT-InCKF) 的状态估计误差较大。当 InCKF 对模型参数的估计稳定后, 二者状态估计误差相应地收敛到较小的范围内。其中, UT-InCKF 状态估计误差大于 SI2-InCKF 算法, 且由图4可知, UT-InCKF 的状态估计误差范数曲线波动较大, 即稳定性差于 SI2-InCKF 算法。由理论分析可知, UT 变换可以达到三阶矩的逼近精度, 对非线性函数的逼近精度应高于 SI2。在 InCKF 算法中, 状态估计由 CKF 算法实现, UT 变换和 SI2 仅使 MAP 从采样假设值中完成对模型参数的选取, 因此可能出现 SI2-InCKF 算法的状态估计结果优于 UT-InCKF 的情况。此外, 图3(b) 显示, SI2 的状态估计精度与模型参数已知情况下 CKF 算法的状态估计精度十分接近, 表明 SI2-InCKF 算法能够有效估计模型的未知参数。

情况2 模型未知参数时变情况下仿真。

设定时变未知参数 $\theta(t)$ 的真值为

$$\theta(t) = \begin{cases} [0.5 \ 25 \ 8 \ 0.05]^T, & 0 \leq t \leq 50; \\ [0.6 \ 30 \ 9.6 \ 0.06]^T, & 50 < t \leq 100. \end{cases} \quad (44)$$

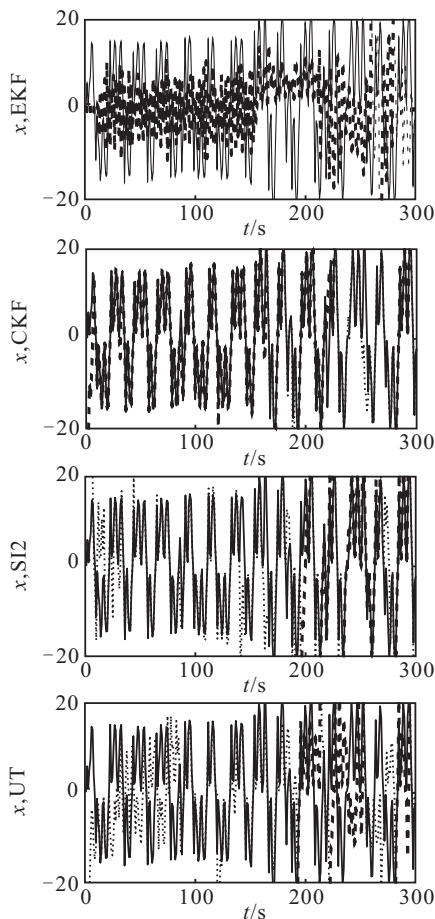


图6 参数 θ 时变情况下的状态真实值和滤波估计结果比较

由式(44)可知, 当时间 $t > 50$ 时, 未知参数真值变为原来的 1.2 倍, 仿真中其他参数的设定同参数设定常情况。

图6~图9分别为参数 θ 时变情况下, 状态真实值和滤波估计结果比较、状态估计误差、状态估计误差范数、模型参数估计误差。由图6、图7和图9可见, 当 $t > 50$, 即模型未知参数真值变为初始设定值的 1.2 倍以后, 扩维 EKF 算法、SI2-InCKF 算法和 UT-InCKF 等算法均出现了不同程度的估计误差。其中, 扩维 EKF 算法的估计偏差最大, 性能表现最差。当 $t = 70$ 以后, SI2-InCKF 算法和 UT-InCKF 算法对系统状态的估计进入较为平稳的阶段, 其中 SI2-InCKF 算法的估计精度最高, 由图8可见其数值稳定性也好于 UT-InCKF 算法。此外, 由图7可见, SI2-InCKF 算法最后阶段的收敛精度与参数已知情况下 CKF 算法的状态估计精度十分接近。由上述分析可知, SI2-InCKF 算法能够处理未知参数时变情况下的状态估计问题。

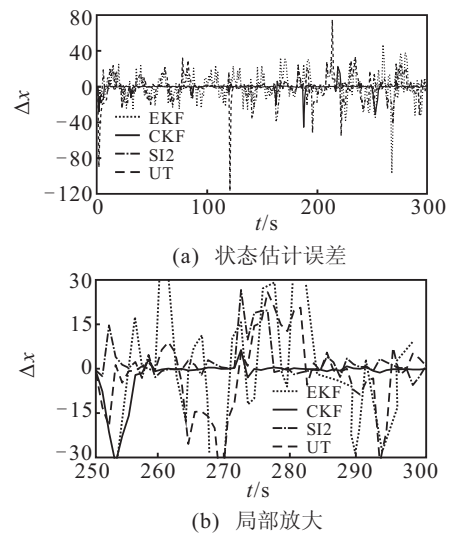


图7 参数 θ 时变情况下的状态估计误差及局部放大

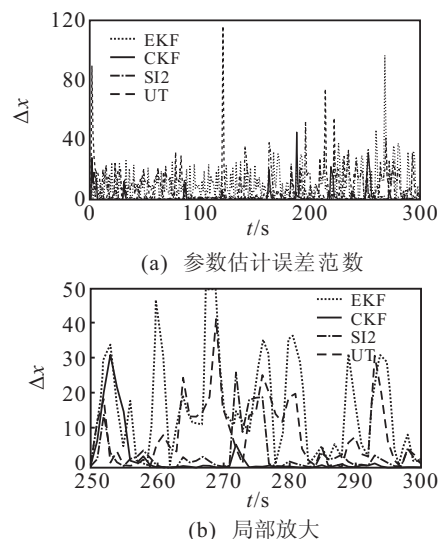


图8 参数 θ 时变情况下的状态估计误差范数及局部放大

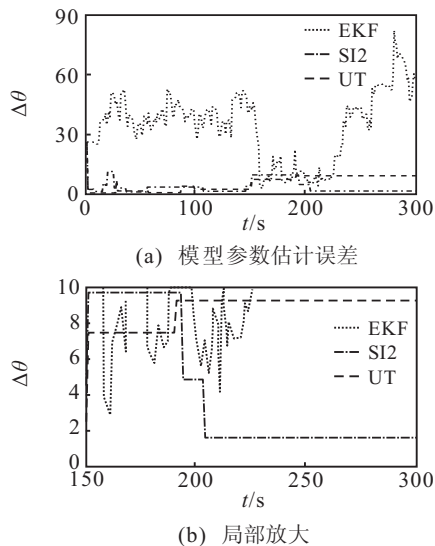


图9 参数 θ 时变情况下的模型参数估计误差及局部放大

综上所述,当非线性系统模型存在定常未知参数或时变未知参数时,扩维EKF算法无法实现对系统状态的准确估计,所提出的InCKF算法能够解决参数未知情况下的状态估计问题,其估计精度和数值稳定性均优于扩维EKF算法.其中,SI2-InCKF算法的精度和数值稳定性高于UT-InCKF,当状态估计过程稳定后,其精度接近于参数已知条件下的CKF算法.

4 结论

本文将极大后验估计与容积卡尔曼滤波算法相结合,提出了一种适用于系统模型参数未知条件下的InCKF算法,并分别基于SI2和UT变换给出概率密度估计两种近似方法.对模型参数样本集进行采样,经过极大后验估计完成对模型参数的识别,从而保证当系统模型参数未知时,InCKF算法对系统的状态估计仍能保持较高的精度,满足应用要求.采用非线性滤波领域的经典测试模型——UNGM进行验证,仿真结果表明,所提出的InCKF算法降低了滤波算法对模型参数精确已知的要求,且具有较高的估计精度,有效地解决了模型参数未知情况下的状态估计问题.

参考文献(References)

- [1] Ghadiok V, Goldin J, Geller D. Gyro-aided vision-based relative pose estimation for autonomous rendezvous and docking[C]. Proc of the 36th Annual AAS Rocky Mountain Section Guidance and Control Conference. Breckenridge, 2013: 713-728.
- [2] Nemra A, Aouf N. Robust INS/GPS sensor fusion for UAV localization using SDRE nonlinear filtering[J]. IEEE Sensors Journal, 2010, 10(4): 789-798.

- [3] Date P, Ponomareva K. Linear and non-linear filtering in mathematical finance: a review[J]. IMA J of Management Mathematics, 2011, 22(3): 195-211.
- [4] Crassidis J L, Junkins J L. Optimal estimation of dynamic systems[M]. Boca Raton: Chapman & Hall/CRC, 2004: 285-292.
- [5] Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation[J]. Proc of the IEEE, 2004, 92(3): 401-422.
- [6] 张勇刚, 黄玉龙, 武哲民, 等. 一种高阶无迹卡尔曼滤波方法[J]. 自动化学报, 2014, 40(5): 838-848.
(Zhang Y G, Huang Y L, Wu Z M, et al. A high order unscented Kalman filtering method[J]. Acta Automatica Sinica, 2014, 40(5): 838-848.)
- [7] 王小旭, 潘泉, 黄鹤, 等. 非线性系统确定采样型滤波算法综述[J]. 控制与决策, 2012, 27(6): 801-812.
(Wang X X, Pan Q, Huang H, et al. Overview of deterministic sampling filtering algorithms for nonlinear system[J]. Control and Decision, 2012, 27(6): 801-812.)
- [8] Gordon N J, Salmond D J, Smith A F M. Novel approach to nonlinear/non-gaussian Bayesian state estimation[J]. IEE Proc of Radar and Signal Processing, 1993, 140(2): 107-113.
- [9] Arasaratnam I, Haykin S. Cubature Kalman filters[J]. IEEE Trans on Automatic Control, 2009, 54(6): 1254-1269.
- [10] Moradkhani H, Sorooshian S, Gupta H V, et al. Dual state-parameter estimation of hydrological models using ensemble Kalman filter[J]. Advances in Water Resources, 2005, 28(2): 135-147.
- [11] Aghili F, Parsa K. Motion and parameter estimation of space objects using laser-vision data[J]. J of Guidance, Control, and Dynamics, 2009, 32(2): 537-549.
- [12] Nrgaard M, Poulsen N K, Ravn O. New developments in state estimation for nonlinear systems[J]. Automatica, 2000, 36(11): 1627-1638.
- [13] Kim S, Park J S. Sequential Monte Carlo filters for abruptly changing state estimation[J]. Probabilistic Engineering Mechanics, 2011, 26(2): 194-201.
- [14] Kim D Y, Jeon M. Square root receding horizon information filters for nonlinear dynamics system models[J]. IEEE Trans on Automatic Control, 2013, 58(5): 1284-1289.
- [15] Fu X Y, Jia Y M. An improvement on resampling algorithm of particle filters[J]. IEEE Trans on Signal Processing, 2010, 58(10): 5414-5420.

(责任编辑: 郑晓蕾)