

学生t量测分布下鲁棒粒子滤波算法的设计与实现

王宗原^{①②} 周卫东^{*①}

^①(哈尔滨工程大学自动化学院 哈尔滨 150001)

^②(哈尔滨工程大学数学科学学院 哈尔滨 150001)

摘要: 野值是一种异于总体数据的非高斯量测值,在实际传输中野值的加入常使信号出现厚尾特性。粒子滤波是基于贝叶斯框架的适用于非线性/非高斯系统的一种滤波方法。如果在量测噪声中存在野值会使粒子滤波的精度下降。该文利用学生t分布建模量测噪声模型,结合变分贝叶斯(VB)递推方法设计一种新颖的边缘粒子滤波(MPF-VBM),它在滤波同时对量测噪声的包括均值在内的全部参数进行实时估计。进一步,利用该估计算法,在量测噪声时变条件下研究了噪声关联的粒子滤波算法(MPF-VBM-COR)。通过对典型单变量增长模型的仿真,验证了所提两种算法相比于已有算法在状态估计上具有更优越的鲁棒性。

关键词: 粒子滤波; 变分贝叶斯; 均值估计; 噪声关联; 学生t分布

中图分类号: TN911.7

文献标识码: A

文章编号: 1009-5896(2019)12-2957-08

DOI: [10.11999/JEIT190144](https://doi.org/10.11999/JEIT190144)

Design and Implementation of Robust Particle Filter Algorithms under Student-t Measurement Distribution

WANG Zongyuan^{①②} ZHOU Weidong^①

^①(College of Automation, Harbin Engineering University, Harbin 150001, China)

^②(College of Mathematical Sciences, Harbin Engineering University, Harbin 150001, China)

Abstract: Outliers are non-Gaussian measurement values far from the bulk of data. In practical transmission, the signals added with outlier often have the heavy-tailed property. Particle filter is based on the Bayesian framework and applicable to the non-linear and non-Gaussian system. However, measurement noise with outlier degrades the performance of particle filter. In this paper, student-t distribution is used to model the measurement noise, combined with Variational Bayes (VB), a novel particle filter Marginalized Particle Filter with VB Mean(MPF-VBM) is designed, which can estimate all parameters of t-distributed measurement distribution including mean parameter as well as state. Further, particle filter with noise correlation (MPF-VBM-COR) at the same epoch which is applicable to time variant measurement noise is developed. For verifying the performances of the proposed algorithms, the simulations on the typical univariate non-stationary growth model are performed under the different noise conditions in detail. The outcomes show that the proposed two algorithms of MPF-VBM and MPF-VBM-COR (MPF-VBM-Correlation) have the superior performances to the compared ones.

Key words: Particle filter; Variational Bayes(VB); Mean estimation; Noise correlation; Student-t distribution

1 引言

粒子滤波是通过对后验滤波密度的量测更新给出离散状态的贝叶斯近似解。不但适用于非高斯非

线性离散系统的状态估计,而且序列重要性重采样是获取样本的主要方法,通过样本与对应似然权值的加权和得到滤波状态^[1]。由于粒子滤波方法结构简单局限性小,因此已被广泛应用到各种实际问题中。

粒子滤波要求系统的模型参数或者量测噪声统计信息保持不变,但在实际应用中可能噪声参数未知且随环境的变化而变化,这时再使用传统粒子滤波算法会使滤波性能下降甚至滤波失效。因此在滤波的同时需要对噪声统计特性或模型结构参数进行更新。针对这个问题,Storvik^[2]给出未知静态参数

收稿日期: 2019-03-13; 改回日期: 2019-07-23; 网络出版: 2019-07-27

*通信作者: 周卫东 zhouweidong@hrbeu.edu.cn

基金项目: 国家自然科学基金(61773133), 中央高校基本科研业务费(3072019CF2419)

Foundation Items: The National Natural Science Foundation of China (61773133), The Fundamental Research Funds of the Central Universities (3072019CF2419)

下的状态和参数同时估计的粒子滤波算法, 通过给定模型和未知噪声参数的充分统计量, 利用已知的参数递推方法更新参数, 得到实时状态估计。文献[3,4]又考虑了统计特性恒定和时变的高斯噪声, 用共轭先验设定Gauss分层参数的分布, 并进行递推更新。变分贝叶斯(Variational Bayes, VB)方法是一种从结构上近似推断的贝叶斯方法, 在假定参数和隐变量独立情况下, 能将联合参数估计问题分解为近似的单变量估计问题, 依据后验密度和真实密度在KL (Kullback-Leibler)距离最大化准则下表示近似程度^[5]。Piche等人^[6]首次提出用VB结合学生t量测分布模型进行状态估计, 从而解决观测过程中出现厚尾非高斯量测, 文献[7,8]用学生t分布分别来建模量测模型和过程模型, 结合VB方法得到新颖的鲁棒卡尔曼滤波。文献[9]将VB方法与粒子滤波结合用来实现对具有学生t分布的厚尾量测模型进行非线性状态估计, 同文献[3,4]一样采用边缘化的方法对状态与参数分别估计, 不同之处在于将学生t分布的自由度参数设为待估计参数, 同尺度参数与隐变量一道, 通过伽马共轭先验分布和VB方法得到参数后验分布表达式。仿真表明, 由于自由度可以调节, 使学生t分布厚尾形状可变, 所以在相同噪声情况下估计性能优于文献[3,4]中所提的方法。粒子滤波与变分贝叶斯方法相结合在文献[10]中也给出一种形式, 但它是利用变分贝叶斯在分割的空间内采样粒子, 以使粒子滤波适用于高维状态空间的估计, 这种形式不同于本文用VB估计参数的研究领域。

从学生t构成形式可以发现, 决定学生t分布对称轴位置的均值变量与尺度参数、自由度参数具有同等重要地位。而且在实际应用中, 如导航^[11]和测距^[4]中, 由于传感器偏移使噪声均值非0且随系统运行发生变化, 所以对均值变量的估计也十分必要。文献[11,12]通过极大后验和VB方法对线性状态模型的量测噪声均值进行了估计。但对于非线性状态空间模型, 使用VB方法对噪声均值的估计还没有涉及。从贝叶斯推断的角度, 均值作为变量与尺度参数相关联, 因此通过VB推导的均值后验分布表达式不同于以往没有均值参与的, 如文献[9]所给形式。在得到量测噪声参数后, 对量测噪声时变的噪声同时关联的粒子滤波问题进行探究。尽管文献[13,14]涉及了粒子滤波的噪声关联问题, 但是它们不能解决量测噪声时变或有野值的情况。

基于以上讨论, 本文用学生t分布对非高斯量测噪声进行建模, 同时考虑均值参数的影响, 将均值与尺度矩阵联合共轭先验为高斯-伽马分布, 利用变分贝叶斯方法结合粒子滤波对时变量测噪声下

非线性状态进行估计得到MPF-VBM算法, 然后通过获得参数估计研究过程噪声依赖时变量测噪声的粒子滤波算法MPF-VBM-COR。文章的其余部分组织如下, 问题公式化在第2节给出, 在学生t量测分布下的参数和状态估计算法及噪声关联算法在第3节给出, 仿真演示的例子在第4节给出, 第5节给出结论。

2 问题公式化

首先考虑如式(1)、式(2)所示的非线性状态空间模型

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

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k \quad (2)$$

其中, $\mathbf{x}_k \in \mathbb{R}^n$ 状态向量, 量测 $\mathbf{z}_k \in \mathbb{R}^m$, $f(\cdot)$ 和 $h(\cdot)$ 都是已知非线性函数, 过程噪声 $\mathbf{w}_k \in \mathbb{R}^n$ 假定服从0均值高斯分布, 具有方差 $E[\mathbf{w}_k \mathbf{w}_k^T] = \mathbf{Q}_k$, 厚尾观测噪声 \mathbf{v}_k 服从对称学生t分布, 分布表达式为

$$S(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Lambda}, \nu) = \frac{\Gamma((\nu + m)/2)}{\Gamma(\nu/2)} \frac{|\boldsymbol{\Lambda}|^{1/2}}{(\pi\nu)^{m/2}} \cdot \left[1 + \frac{1}{\nu} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Lambda} (\mathbf{x} - \boldsymbol{\mu}) \right]^{-(\nu+m)/2} \quad (3)$$

其中, $S(\cdot)$ 表示学生t分布, $\Gamma(\cdot)$ 表示伽马函数, $\boldsymbol{\mu}$, $\boldsymbol{\Lambda}$ 和 ν 分别表示均值, 尺度矩阵和自由度参数。应该指出的是学生t分布是一种厚尾的高斯分布, 当 $\nu \rightarrow \infty$ 时就成为具有相同 $\boldsymbol{\mu}$ 的对应高斯分布。从式(3)可见, 即使对于同一个数据点 \mathbf{x} , 均值的不同也会使密度值不同。所以准确知道均值变量十分重要。为进行利用VB的参数估计, 不失一般性, 假定 \mathbf{v}_k 的尺度矩阵 $\boldsymbol{\Lambda}_k$ 为对角矩阵, 即 $\boldsymbol{\Lambda}_k = \text{diag}(\Lambda_{k,1}, \Lambda_{k,2}, \dots, \Lambda_{k,m})$ 。对于式(1)、式(2)所给状态空间模型, 依据学生t的量测噪声分布特点, 时刻 k 的量测似然可表示为具有隐变量 u_k 的如式(4)、式(5)的高斯分层形式

$$p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \boldsymbol{\mu}_k, u_k, \boldsymbol{\Lambda}_k) = N(\mathbf{z}_k | h(\mathbf{x}_k^{(i)}) + \boldsymbol{\mu}_k, (u_k \boldsymbol{\Lambda}_k)^{-1}) \quad (4)$$

$$p(u_k | \nu_k) = G(u_k | \nu_k/2, \nu_k/2) \quad (5)$$

其中, u_k 是引入的隐变量, $G(\cdot | \alpha, \beta)$ 表示服从参数为 α, β 的伽马分布。由文献[15]所示, 均值和尺度矩阵的联合共轭先验服从高斯-伽马分布, 并且可分解为各自单独分布的乘积

$$p(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) = p(\boldsymbol{\mu}_k | \boldsymbol{\Lambda}_k) p(\boldsymbol{\Lambda}_k) = \prod_{j=1}^m \left[N(\boldsymbol{\mu}_{k,j} | \boldsymbol{\eta}_{k|k-1,j}, (\beta_{k|k-1} \Lambda_{k,j})^{-1}) \cdot G(\Lambda_{k,j} | c_{k|k-1,j}, d_{k|k-1,j}) \right] \quad (6)$$

其中， $\boldsymbol{\eta}_{k|k-1,j}$ 和 $(\beta_{k|k-1}\Lambda_{k,j})^{-1}$ 分别为 $\boldsymbol{\mu}_{k,j}$ 的均值和方差， $\beta_{k|k-1}$ 为精度系数。

自由度 ν_k 的共轭先验服从伽马分布

$$p(\nu_k) = G(\nu_k | a_{k|k-1}, b_{k|k-1}) \quad (7)$$

为了更清晰地说明各变量之间的关系，用图1所示的有向图对 k 时刻状态，量测和超参数的贝叶斯分层模型进行表示：

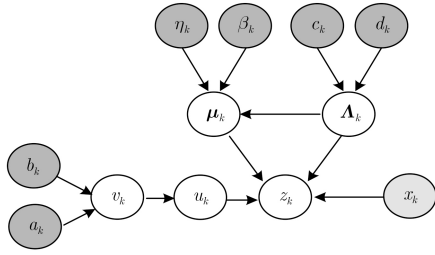


图1 贝叶斯推断图模型

从图1中可见，相比于在文献[9]，两个额外的超参数 β_k 和 $\boldsymbol{\eta}_k$ 被加入进来，并且 $\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k$ 之间具有关联性，相比于文献[6]，自由度变量具有超参数 a_k, b_k 。

3 主要结果

本节主要应用VB对式(1)–式(7)的中参数的变分后验进行推导。VB方法的原理及公式详见文献[7,15]。

首先，状态、量测和参数的联合分布可分解为

$$\begin{aligned} p(\mathbf{x}_k^{(i)}, \boldsymbol{\mu}_k, u_k, \boldsymbol{\Lambda}_k, \nu_k | \mathbf{z}_{1:k}) &\propto p(\mathbf{x}_k^{(i)}, | \mathbf{z}_{1:k-1}) \\ &\cdot p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \boldsymbol{\mu}_k, u_k, \boldsymbol{\Lambda}_k) p(\boldsymbol{\mu}_k | \boldsymbol{\Lambda}_k) p(\boldsymbol{\Lambda}_k) \\ &\cdot p(u_k | \nu_k) p(\nu_k) \end{aligned} \quad (8)$$

其中， $\mathbf{x}_k^{(i)}$ 表示 k 时刻第 i 个粒子，令 $\boldsymbol{\Theta}_k = \{\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k, \nu_k\}$ 表示参数集，利用边缘化思想可分别进行状态与参数的估计，在时刻 k ，状态参数联合后验分布为

$$\begin{aligned} p(\mathbf{x}_k, \boldsymbol{\Theta}_k | \mathbf{z}_{1:k}) &= p(\boldsymbol{\Theta}_k | \mathbf{x}_k, \mathbf{z}_{1:k}) p(\mathbf{x}_k | \mathbf{z}_{1:k}) \\ &= \sum_{i=1}^N \omega_k^{(i)} p(\boldsymbol{\Theta}_k^{(i)} | \mathbf{x}_k^{(i)}, \mathbf{z}_{1:k}) \\ &\quad \cdot \delta(\mathbf{x}_k - \mathbf{x}_k^{(i)}) \end{aligned} \quad (9)$$

而且

$$\omega_k^{(i)} = \omega_{k-1}^{(i)} p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \boldsymbol{\Theta}_k^{(i)}) \quad (10)$$

其中， $\delta(\cdot)$ 表示Dirac-Delta函数。 $\omega_k^{(i)}$ 是对应第 i 个粒子的重要性权值，式(10)中的似然 $p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \boldsymbol{\Theta}_k^{(i)})$ 服从学生t分布，可见它不但依赖粒子样本值，而

且依赖于 $k-1$ 时刻参数一步预测的样本值。下面给出在量测更新后， $\boldsymbol{\Theta}_k^{(i)}$ 依据VB的具体更新过程。

首先，对于均值变量 $\boldsymbol{\mu}_k$ 和尺度矩阵 $\boldsymbol{\Lambda}_k$ ，通过基本贝叶斯公式 $p(x|y) = p(x, y)/p(y)$ ， $(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k)$ 联合变分后验的对数函数表示为

$$\begin{aligned} \ln q(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) &= E_{u_k, \nu_k} \left[\ln p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \boldsymbol{\mu}_k, u_k, \boldsymbol{\Lambda}_k) \right. \\ &\quad \left. + \ln p(\boldsymbol{\mu}_k | \boldsymbol{\Lambda}_k) + \ln p(\boldsymbol{\Lambda}_k) \right] + C_{\boldsymbol{\mu}, \boldsymbol{\Lambda}} \\ &= \frac{1}{2} \sum_{j=1}^m \ln \Lambda_{k,j} - \frac{1}{2} \beta_{k|k-1} \\ &\quad \cdot \sum_{j=1}^m \Lambda_{k,j} (\boldsymbol{\mu}_k - \boldsymbol{\eta}_{k|k-1})_j^2 \\ &\quad + \frac{m}{2} E[\ln u_k] - \frac{1}{2} E[u_k] \\ &\quad \cdot \sum_{j=1}^m \Lambda_{k,j} (\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k)_j^2 \\ &\quad + \sum_{j=1}^m [(c_{k|k-1,j} - 1) \ln \Lambda_{k,j} \\ &\quad - d_{k|k-1,j} \Lambda_{k,j}] + C_{\boldsymbol{\mu}, \boldsymbol{\Lambda}} \end{aligned} \quad (11)$$

其中， $E_x[\cdot]$ 表示括号中的概率分布对变量 x 求期望， $C_{\boldsymbol{\mu}, \boldsymbol{\Lambda}}$ 是独立于 $\boldsymbol{\mu}$ 和 $\boldsymbol{\Lambda}$ 的常数项， $(\cdot)_j$ 表示括号中向量的第 j 个分量。 $\boldsymbol{\mu}_k$ 边缘后验对数函数为

$$\begin{aligned} \ln q(\boldsymbol{\mu}_k | \boldsymbol{\Lambda}_k) &= E_{u_k, \nu_k} \left[\ln p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \boldsymbol{\mu}_k, u_k, \boldsymbol{\Lambda}_k) \right. \\ &\quad \left. + \ln p(\boldsymbol{\mu}_k | \boldsymbol{\Lambda}_k) \right] + C \\ &= -\frac{1}{2} E[u_k] \sum_{j=1}^m \Lambda_{k,j} \\ &\quad \cdot (\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k)_j^2 \\ &\quad - \frac{1}{2} \beta_{k|k-1} \sum_{j=1}^m \Lambda_{k,j} (\boldsymbol{\mu}_k - \boldsymbol{\eta}_{k|k-1})_j^2 \\ &\quad + C_{\boldsymbol{\mu}} \end{aligned} \quad (12)$$

由于式(12)是 $\boldsymbol{\mu}_k$ 的2次式，所以变分后验 $q(\boldsymbol{\mu}_k | \boldsymbol{\Lambda}_k)$ 服从正态分布 $N(\boldsymbol{\mu}_k | \boldsymbol{\eta}_{k|k}, (\beta_{k|k} \boldsymbol{\Lambda}_{k|k})^{-1})$ ，其中参数的递推公式为

$$\beta_{k|k} = \beta_{k|k-1} + E[u_k] \quad (13)$$

$$\begin{aligned} \boldsymbol{\eta}_{k|k} &= \boldsymbol{\eta}_{k|k-1} + E[u_k] (\beta_{k|k})^{-1} \\ &\quad \cdot (\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\eta}_{k|k-1}) \end{aligned} \quad (14)$$

$\boldsymbol{\Lambda}_k$ 的变分后验分布对数式由式(11)减式(12)得到

$$\begin{aligned} \ln q(\Lambda_k) = & -\frac{1}{2} \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right)^T \\ & \cdot \mathbf{E}[u_k] \Lambda_k \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right) \\ & - \frac{1}{2} \left(\boldsymbol{\mu}_k - \boldsymbol{\eta}_{k|k-1} \right)^T \beta_{k|k-1} \Lambda_k \left(\boldsymbol{\mu}_k - \boldsymbol{\eta}_{k|k-1} \right) \\ & + \sum_{j=1}^m \left[(c_{k|k-1,j} - 1) \ln \Lambda_{k,j} - d_{k|k-1,j} \Lambda_{k,j} \right] \\ & + \frac{1}{2} \left(\boldsymbol{\mu}_k - \boldsymbol{\eta}_{k|k} \right)^T (\beta_{k|k} \Lambda_k) \\ & \cdot \left(\boldsymbol{\mu}_k - \boldsymbol{\eta}_{k|k} \right) + C \end{aligned} \quad (15)$$

将式(13), 式(14)代入式(15), 通过表达式化简和共轭先验的概念, 得到 $\Lambda_{k,j}$ 变分后验服从伽马分布 $q(\Lambda_{k,j}) = G(\Lambda_{k,j} | c_{k|k,j}, d_{k|k,j})$, 其中参数

$$c_{k|k,j} = c_{k|k-1,j} + \frac{1}{2} \quad (16)$$

$$\begin{aligned} d_{k|k,j} = & d_{k|k-1,j} + \frac{1}{2} \mathbf{E}[u_k] \left[1 - \mathbf{E}[u_k] (\beta_{k|k})^{-1} \right] \\ & \cdot \left[\left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\eta}_{k|k-1} \right) \right. \\ & \cdot \left. \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\eta}_{k|k-1} \right)^T \right]_{jj} \end{aligned} \quad (17)$$

其中, $[\mathbf{A}]_{jj}$ 表示矩阵 \mathbf{A} 主对角线第 jj 个元素。

隐变量 u_k 变分后验 $q(u_k)$ 的对数函数表示为

$$\begin{aligned} \ln q(u_k) = & \mathbf{E}_{\boldsymbol{\mu}_k, \Lambda_k, \nu_k} \left[\ln p \left(\mathbf{z}_k \mid \mathbf{x}_k^{(i)}, \boldsymbol{\mu}_k, u_k, \Lambda_k \right) \right. \\ & \left. + \ln p(u_k | \nu_k) \right] + C_u \\ = & \frac{\mathbf{E}[\nu_k] - 1}{2} \ln u_k - \frac{\mathbf{E}[\nu_k]}{2} u_k + \frac{m}{2} \ln u_k \\ & + \frac{1}{2} \ln |\Lambda_k| - \frac{1}{2} \mathbf{E}_{\boldsymbol{\mu}_k, \Lambda_k} \left[\left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) \right. \right. \\ & \left. \left. - \boldsymbol{\mu}_k \right)^T (u_k \Lambda_k) \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right) \right] \\ & + C_u \end{aligned} \quad (18)$$

对式(18)指数化运算, $q(u_k)$ 服从伽马分布 $q(u_k) = G(u_k | \nu_k^1, \nu_k^2)$, 其中

$$\begin{aligned} \nu_k^1 = & \frac{1}{2} (\mathbf{E}[\nu_k] + m) \\ \nu_k^2 = & \frac{1}{2} \left(\mathbf{E}[\nu_k] + \mathbf{E}_{\boldsymbol{\mu}_k, \Lambda_k} \left[\left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right)^T \right. \right. \\ & \cdot \Lambda_k \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right) \left. \left. \right] \right) \\ = & \frac{1}{2} \left[\mathbf{E}[\nu_k] + \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\eta}_{k|k} \right)^T \right. \\ & \cdot \mathbf{E}[\Lambda_k] \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\eta}_{k|k} \right) + m \beta_{k|k}^{-1} \left. \right] \end{aligned} \quad (19)$$

这里, 在 $\mathbf{E}_{\boldsymbol{\mu}_k, \Lambda_k} \left[\left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right)^T \Lambda_k \left(\mathbf{z}_k - h(\mathbf{x}_k^{(i)}) - \boldsymbol{\mu}_k \right) \right]$ 推导中, 使用了 $\boldsymbol{\mu}_k$ 的如式(20)的2阶矩

$$\mathbf{E}_{\boldsymbol{\mu}_k, \Lambda_k} [\boldsymbol{\mu}_k \boldsymbol{\mu}_k^T] - \mathbf{E}[\boldsymbol{\mu}_k] \mathbf{E}[\boldsymbol{\mu}_k]^T = (\beta_{k|k} \mathbf{E}[\Lambda_k])^{-1} \quad (20)$$

自由度 ν_k 的近似后验为

$$\begin{aligned} \ln q(\nu_k) = & \frac{\nu_k}{2} \ln \frac{\nu_k}{2} - \ln \Gamma\left(\frac{\nu_k}{2}\right) + \left(\frac{\nu_k}{2} - 1\right) \\ & \cdot \mathbf{E}[\ln u_k] - \frac{\nu_k}{2} \mathbf{E}[u_k] + (a_{k|k-1} - 1) \ln \nu_k \\ & - b_{k|k-1} \nu_k + C_v \end{aligned} \quad (21)$$

使用Stirling的近似 $\ln \Gamma\left(\frac{\nu_k}{2}\right) \approx \frac{\nu_k - 1}{2} \ln \frac{\nu_k}{2} - \frac{\nu_k}{2}$, $q(\nu_k)$ 为伽马分布 $q(\nu_k) = G(\nu_k | a_{k|k}, b_{k|k})$, 其中

$$a_{k|k} = a_{k|k-1} + \frac{1}{2} \quad (22)$$

$$b_{k|k} = b_{k|k-1} + \frac{1}{2} \mathbf{E}[u_k] - \frac{1}{2} \mathbf{E}[\ln u_k] - \frac{1}{2} \quad (23)$$

在以上每一个参数推导过程中, 都需要对其余参数进行期望操作, 根据分布期望公式, 参数期望如式(24)给出

$$\left. \begin{aligned} \mathbf{E}[u_k] = & \nu_k^1 / \nu_k^2, \quad \mathbf{E}[\ln u_k] = \Psi(\nu_k^1) - \ln \nu_k^2 \\ \mathbf{E}[\Lambda_{k,t}] = & c_{k|k,t} / d_{k|k,t}, \quad \mathbf{E}[\nu_k] = a_{k|k} / b_{k|k}, \\ \mathbf{E}[\boldsymbol{\mu}_k] = & \boldsymbol{\eta}_{k|k} \end{aligned} \right\} \quad (24)$$

其中, $\Psi(\cdot)$ 为di-gamma函数。

尽管以上参数更新相互耦合, 但这一问题解决可通过对任一参数初始化实现, 然后依次迭代更新。同时可见, 如果均值变量为0, 涉及均值的超参数 η_k 和 β_k 的递推式(13)、式(14)以及在式(17)和式(19)含有的 η_k 和 β_k 都消失了, 算法退化为文献[9]所提算法。因此所提算法可视为文献[9]算法的丰富和扩展。

噪声可能为时变的情形, 因此引入遗忘因子 $\rho \in (0, 1)$ 标识噪声波动程度, 噪声参数时间更新模型为

$$\left. \begin{aligned} \boldsymbol{\eta}_{k|k-1} = & \rho \boldsymbol{\eta}_{k-1|k-1}, \quad \beta_{k|k-1} = \rho \beta_{k-1|k-1}, \\ a_{k|k-1} = & \rho a_{k-1|k-1}, \quad b_{k|k-1} = \rho b_{k-1|k-1}, \\ c_{k|k-1,t} = & \rho c_{k-1|k-1,t}, \quad d_{k|k-1,t} = \rho d_{k-1|k-1,t} \end{aligned} \right\} \quad (25)$$

表1给出所提算法1的流程。

依据表1, 接着考虑过程噪声依赖时变量测噪声的粒子滤波算法。将学生t量测似然式(4)改写成高斯分层形式

$$p(\mathbf{z}_k | \mathbf{x}_k^{(i)}, \mathbf{R}_k, u_k) = N\left(\mathbf{z}_k \mid h(\mathbf{x}_k^{(i)}) + \boldsymbol{\mu}_k, \frac{\mathbf{R}_k}{u_k}\right) \quad (26)$$

表1 基于VB带有噪声均值估计的边缘粒子滤波(MPF-VBM)算法

<p>从分布 $N(\mathbf{x}_0 \mathbf{m}_{0 0}, \mathbf{P}_{0 0})$ 采样粒子 $\mathbf{x}_0^{(i)} i = 1, 2, \dots, N$, 并且设置权重 $\omega_0^{(i)} = 1/N$; 初始化超参数 $a_0^{(i)}, b_0^{(i)}, c_0^{(i)}, d_0^{(i)}, \nu_0^{(i)}, \eta_0^{(i)}$ 和 $\beta_0^{(i)}$; 计算初始参数 μ_0, Λ_0 和 ν_0 期望。对时刻 $k = 1, 2, \dots, K$</p> <p>对每一个粒子 $i = 1, 2, \dots, N$</p> <p>(1) 使用式(25)做噪声参数的时间更新;</p> <p>(2) 从状态传递方程 $p(\mathbf{x}_k^{(i)} \mathbf{x}_{k-1}^{(i)})$ 做粒子的一步预测;</p> <p>(3) 通过新量测 z_k 用式(15)更新重要性权重;</p> <p>(4) 必要的话, 粒子重采样;</p> <p>(5) 使用式(13), 式(14), 式(16), 式(17), 式(18), 式(19), 式(22), 式(23), 并利用重采样粒子做噪声参数后验更新,</p> <p>$\Theta_{k k}^{(i)} = T(\Theta_{k k-1}^{(i)}, x_k^{(i)}, z_k)$; 其中 $T(\cdot)$ 代表参数充分统计量;</p> <p>得出当前的噪声参数期望值及对应状态值;</p> <p>在 $k == K$ 前, 进行下一次循环。</p>
--

$$p(u_k | \nu_k) = G(u_k | \nu_k/2, \nu_k/2) \quad (27)$$

其中, 正态分布的方差表达式为 \mathbf{R}_k/u_k , \mathbf{R}_k 表示式(4)中 $(\Lambda_k)^{-1}$ 。假定过程噪声依赖量测噪声的关联矩阵为 \mathbf{S}_k , 即

$$\mathbf{E}[\mathbf{w}_k \mathbf{v}_k^T] = \mathbf{S}_k \quad (28)$$

同时假定 $\mathbf{w}_k, \mathbf{v}_k$ 各自单独序列仍然是相互独立的。下面利用解关联框架进行噪声解关联操作, 根据式(2), $\mathbf{z}_k - h(\mathbf{x}_k) - \mathbf{v}_k = 0$, 那么状态传递方程式(1)修改为

$$\begin{aligned} \mathbf{x}_{k+1} &= f(\mathbf{x}_k) + \mathbf{w}_k + \mathbf{J}_k(\mathbf{z}_k - h(\mathbf{x}_k) - \mathbf{v}_k) \\ &= f^*(\mathbf{x}_k) + \mathbf{J}_k \mathbf{z}_k + \mathbf{w}_k^* \end{aligned} \quad (29)$$

这里, $f^*(\mathbf{x}_k) = f(\mathbf{x}_k) - \mathbf{J}_k h(\mathbf{x}_k)$, $\mathbf{w}_k^* = \mathbf{w}_k - \mathbf{J}_k \mathbf{v}_k$, \mathbf{J}_k 是要确定的辅助参数。 \mathbf{w}_k^* 可被看成新的过程噪声, 由于是高斯分布的线性组合, 所以仍然服从高斯分布, 且它与量测噪声 \mathbf{v}_k 的协方差为

$$\begin{aligned} \mathbf{E}[\mathbf{w}_k^* \mathbf{v}_k^T] &= \mathbf{E}[\mathbf{w}_k \mathbf{v}_k] - \mathbf{J}_k \mathbf{E}[\mathbf{v}_k \mathbf{v}_k^T] \\ &= \mathbf{S}_k - \mathbf{J}_k (\mathbf{E}[\mathbf{R}_k] / \mathbf{E}[u_k]) \end{aligned} \quad (30)$$

令协方差 $\mathbf{E}[\mathbf{w}_k^* \mathbf{v}_k^T] = 0$, 得到辅助矩阵 \mathbf{J}_k 为

$$\mathbf{J}_k = \mathbf{S}_k \left(\mathbf{E}[u_k] (\mathbf{E}[\mathbf{R}_k])^{-1} \right) \quad (31)$$

式(29), 式(2)连同式(31)一起给出噪声同时关联的等价状态空间模型。通过式(31), 等价传递模型的高斯过程噪声 \mathbf{w}_k^* 的均值与方差为

$$\begin{aligned} \mathbf{E}[\mathbf{w}_k^*] &= \boldsymbol{\varepsilon}_k^* = \mathbf{E}[\mathbf{w}_k] - \mathbf{J}_k \mathbf{E}[\mathbf{v}_k] \\ &= -\mathbf{S}_k \left(\mathbf{E}[u_k] (\mathbf{E}[\mathbf{R}_k])^{-1} \right) \boldsymbol{\mu}_k \end{aligned} \quad (32)$$

$$\begin{aligned} \text{var}[\mathbf{w}_k^*] &= \mathbf{Q}_k^* = \mathbf{Q}_k + \mathbf{J}_k (\mathbf{E}[\mathbf{R}_k] / \mathbf{E}[u_k]) \mathbf{J}_k^T \\ &\quad - \mathbf{S}_k \mathbf{J}_k^T - \mathbf{J}_k \mathbf{S}_k^T \\ &= \mathbf{Q}_k - \mathbf{S}_k \left(\mathbf{E}[u_k] (\mathbf{E}[\mathbf{R}_k])^{-1} \right) \mathbf{S}_k^T \end{aligned} \quad (33)$$

因此, 由式(1), 式(2)组成具有噪声关联的系

统转化为由式(29), 式(2)组成不含噪声关联的等价系统滤波问题。算法流程与表1算法相同, 就是在第(2)步中增加对 \mathbf{w}_k^* 噪声参数的计算式(32)和式(33), 状态转移模型由 $p_{\mathbf{w}_k}(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)})$ 替换为 $p_{\mathbf{w}_k^*}(\mathbf{x}_k^{(i)} | \mathbf{x}_{k-1}^{(i)}, \mathbf{z}_{k-1})$, 这样得到时变量测噪声下噪声同时关联的粒子滤波算法, 记作MPF-VB-COR-1(Marginalized Particle Filter with VB for Correlation-1)。

4 仿真实验

这里考虑在文献[4,9]广泛使用的如下非线性模型来验证上面所提两种算法的性能

$$\left. \begin{aligned} x_k &= 0.5x_{k-1} + 25 \frac{x_{k-1}}{1 + x_{k-1}^2} \\ &\quad + 8 \cos(1.2(k-1)) + w_{k-1} \\ z_k &= 0.05x_k^2 + v_k \end{aligned} \right\} \quad (34)$$

算法1 初始状态 $\mathbf{x}_0 \sim N(0, 5)$, 过程噪声分布 $p(w_k) = N(w_k | 0, 5)$, 量测噪声 v_k 服从学生t分布, 初始超参数为: $a_0 = b_0 = 0.12, c_0 = 2, d_0 = 5, \eta_0 = 1$ 和 $\beta_0 = 2$, 遗忘因子选取根据文献[9]的遗忘因子与超参数收敛时间的比较说明, 设为 $\rho = 1 - \exp(-4)$ 。仿真时间长度设为1000。根据文献[9]的粒子数与状态均方根误差平均值(ARMSE)的关系, 考虑程序运行的时间因素, 粒子数设为100。Monte Carlo仿真运行30次, 所提的MPF-VBM算法与另外两种算法MPF-VBS (Student)^[9]和MPF-CP (Conjugate Prior)^[3]在相同仿真状况下进行均方根误差(RMSE)^[8]及(ARMSE)^[9]比较。仿真设置3种不同情况噪声, 噪声及对应的状态ARMSE及时间如表2所示。

$U(\cdot)$ 表示均匀分布, 野值以随机形式进入噪声分布。第1种标准高斯噪声加入20%服从均匀分布野值, 噪声均值为0; 第2种高斯噪声无野值加入, 均值为6; 第3种高斯噪声均值为6有20% $U(20, 60)$

表 2 对应3种噪声的3种算法的均方根误差平均值(ARMSE)

量测噪声	MPF-VBM	MPF-VBS	MPF-CP
N(0,1)+20%U(-20, 20)野值	5.5673	5.5468	8.5571
N(6, 1)无野值	4.6170	8.1141	8.8132
N(6,5)+20%U(20, 60)野值	6.3353	8.6288	8.7588
运行时间(s)	0.01067	0.007187	0.005587

的野值加入。MPF-VBM与MPF-CP量测噪声均值估计比较图如图2所示。

从图2可见，本文MPF-VBM算法对3种噪声情况的均值都具有稳健而准确的估计能力，相比较而言，MPF-CP的均值估计易受到野值干扰，跟踪真值时的震荡性较强。但从图2(b)估计图所示，在没有野值干扰时，MPF-CP均值估计具有收敛速度快优点。

图3—图5给出了3种算法对应3种噪声的RMSE的比较图，为了清晰起见，这里只列出了300~330 s时间范围内的RMSE的比较曲线。从图3—图5可见，所提MPF-VBM算法对3种噪声都始终处于估计性能最优越的算法行列。3种算法对应3种噪声ARMSE如表2所示，结果验证了RMSE图的显示。在运行时间上，MPF-VBM近似为MPF-VBS时间的1.5倍，为MPF-CP时间的2倍。量级为ms级。

算法2 所提噪声关联的MPF-VB-COR与算法MPF-VBM及文献[13]中的算法PF-COR在相同噪声关联情况下进行性能比较。Monte Carlo仿真次数为30次。过程噪声与量测噪声超参数设置与算法1相同。设过程噪声 w_k 量测噪声 v_k 有关系式 $w_k = T_k v_k$ ，其中 T_k 为常矩阵，此处设为 $[T_k] = 0.8$ ，得到关联矩

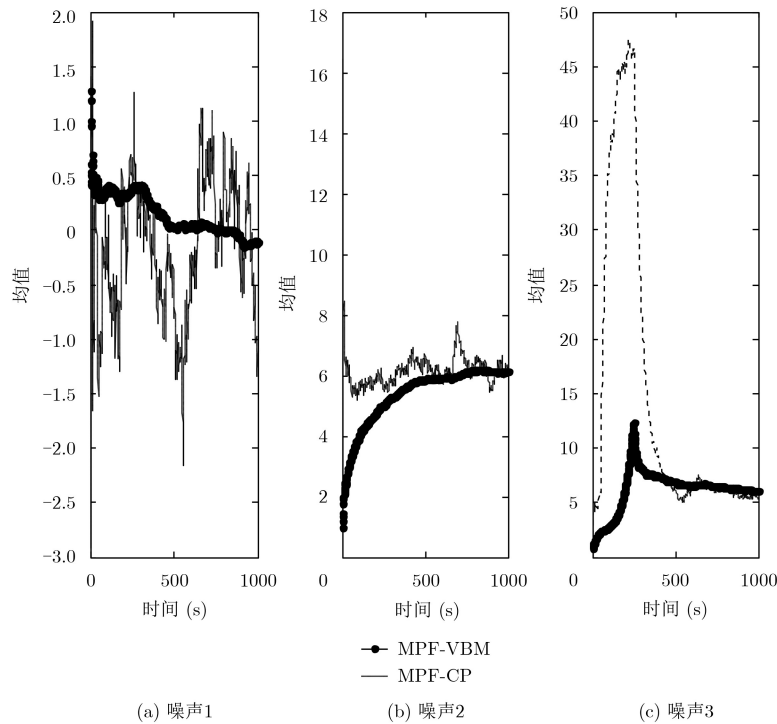


图 2 算法噪声均值估计比较图

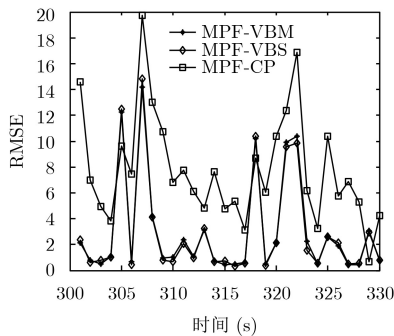


图 3 3种算法对应第1种噪声RMSE比较图

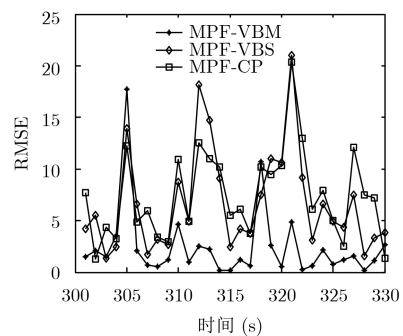


图 4 3种算法对应第2种噪声RMSE比较图

阵 $S_k = E[w_k v_k^T] = T_k E[v_k v_k^T] = T_k E[R_k] / E[u_k]$ 。为了更清晰表示算法性能，图6—图8给出单次Monte Carlo状态的ARMSE的曲线比较图，表3给出设置的量测噪声及30次Monte Carlo的ARMSE。其中，第2种噪声方差R如下变化，当 $t \in [1\text{ s} : 300\text{ s}]$, $R = 8$ ；当 $t \in [301\text{ s} : 700\text{ s}]$, $R = 20$ ；当 $t \in [701\text{ s} : 1000\text{ s}]$, $R = 2$ 。

从图6—图8及表3可见，当第1种量测噪声是参数确定的高斯噪声时，PF-COR的估计性能最好，本文的MPF-VB-COR-1与PF-COR估计性能相近，两者都好于没有考虑关联性的MPF-VBM，从ARMSE曲线看，含有变分估计方法的估计一致性

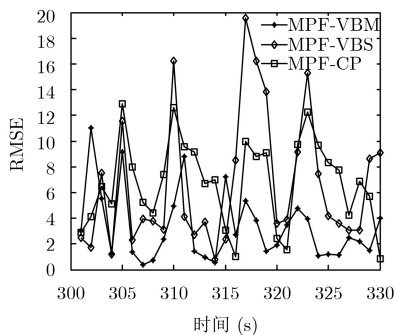


图5 3种算法对应第3种噪声RMSE比较图

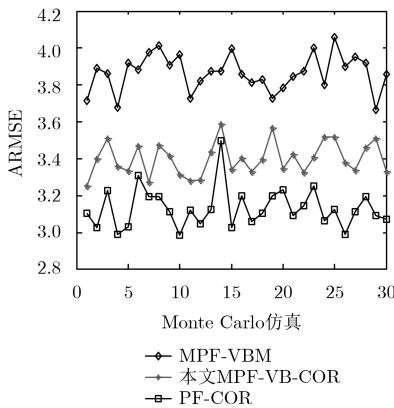


图6 3种算法对第1种关联噪声的ARMSE比较图

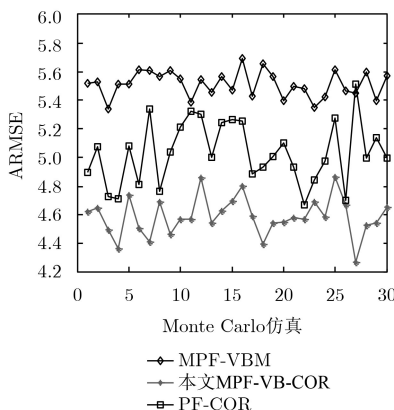


图7 3种算法对第2种关联噪声的状态ARMSE比较图

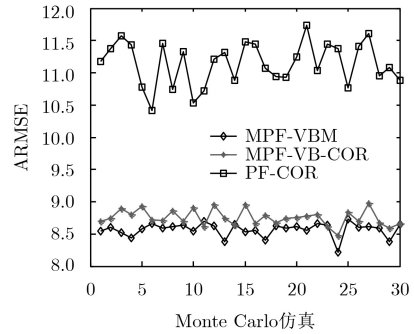


图8 3种算法对第3种关联噪声的状态ARMSE比较图

表3 对应3种噪声的3种算法的均方根误差平均值(ARMSE)

量测噪声	MPF-VBM	MPF-VB-COR	PF-COR
$N(0, 5/2)$	3.8662	3.4025	3.1384
$N(0, R)$ R时变	5.5178	4.5536	5.0392
$N(0, 5/2)+20\%$	8.5892	8.7574	11.0393
Unif(-5, 5)野值			

更好；第2种量测噪声是方差时变的高斯噪声，MPF-VB-COR的估计性能好于PF-COR，两者的性能都优于MPF-VBM，仍然说明考虑了噪声关联情况粒子滤波算法此时更好。在第3种具有野值的高斯噪声，MPF-VB-COR与MPF-VBM的性能基本相同，好于PF-COR，说明此时具有抑制野值干扰的滤波算法具有更好性能，而所提MPF-VB-COR也是具有这一能力的。

5 结论

本文讨论了具有学生t量测噪声的鲁棒粒子滤波算法。量测噪声被设计成服从学生t分布，利用变分贝叶斯方法对t分布的噪声包括均值在内的全部参数进行了估计。先给出了参数的共轭先验，然后推导出参数变分后验表达式。通过边缘化思想给出估计参数值对应的状态值。将此算法与已有学生t量测噪声的粒子滤波方法比较，对于给出的噪声情况，所提粒子滤波展现出其有效性和优越的性能。接着基于所提鲁棒粒子滤波框架，探究了量测噪声时变的噪声关联的粒子滤波，仿真表明当量测噪声时变或有野值存在时，所提噪声关联的粒子滤波算法优于已有噪声关联粒子滤波算法。所提算法运行时间一般为 10^{-2} s的量级。

参考文献

[1] XU Long, MA Kemao, LI Wenshuo, et al. Particle filtering for networked nonlinear systems subject to random one-step sensor delay and missing measurements[J]. *Neurocomputing*, 2018, 275: 2162–2169. doi: 10.1016/j.neucom.2017.10.059.

[2] STORVIK G. Particle filters for state-space models with the presence of unknown static parameters[J]. *IEEE*

- Transactions on Signal Processing*, 2002, 50(2): 281–289. doi: [10.1109/78.978383](https://doi.org/10.1109/78.978383).
- [3] SAHA S, ÖZKAN E, GUSTAFSSON F, *et al.* Marginalized particle filters for Bayesian estimation of Gaussian noise parameters[C]. The 13th International Conference on Information Fusion, Edinburgh, UK, 2010: 1–8. doi: [10.1109/ICIF.2010.5712016](https://doi.org/10.1109/ICIF.2010.5712016).
- [4] ÖZKAN E, ŠMÍDL V, SAHA S, *et al.* Marginalized adaptive particle filtering for nonlinear models with unknown time-varying noise parameters[J]. *Automatica*, 2013, 49(6): 1566–1575. doi: [10.1016/j.automatica.2013.02.046](https://doi.org/10.1016/j.automatica.2013.02.046).
- [5] ZHAO Yujia, FATEHI A, and HUANG Biao. Robust estimation of ARX models with time varying time delays using variational Bayesian approach[J]. *IEEE Transactions on Cybernetics*, 2018, 48(2): 532–542. doi: [10.1109/TCYB.2016.2646059](https://doi.org/10.1109/TCYB.2016.2646059).
- [6] PICHE R, SÄRKKÄ S, and HARTIKAINEN J. Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate student-t distribution[C]. 2012 IEEE International Workshop on Machine Learning for Signal Processing, Santander, Spain, 2012: 1–6. doi: [10.1109/MLSP.2012.6349794](https://doi.org/10.1109/MLSP.2012.6349794).
- [7] ZHANG Yonggang, JIA Guangle, LI Ning, *et al.* A novel adaptive Kalman filter with colored measurement noise[J]. *IEEE Access*, 2018, 6: 74569–74578. doi: [10.1109/ACCESS.2018.2883040](https://doi.org/10.1109/ACCESS.2018.2883040).
- [8] HUANG Yulong, ZHANG Yonggang, LI Ning, *et al.* A novel robust Student's t-based Kalman filter[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017, 53(3): 1545–1554. doi: [10.1109/TAES.2017.2651684](https://doi.org/10.1109/TAES.2017.2651684).
- [9] XU Dingjie, SHEN Chen, and SHEN Feng. A robust particle filtering algorithm with non-Gaussian measurement noise using student-t distribution[J]. *IEEE Signal Processing Letters*, 2014, 21(1): 30–34. doi: [10.1109/LSP.2013.2289975](https://doi.org/10.1109/LSP.2013.2289975).
- [10] AIT-EL-FQUIH B and HOTEIT I. A variational Bayesian multiple particle filtering scheme for large-dimensional systems[J]. *IEEE Transactions on Signal Processing*, 2016, 64(20): 5409–5422. doi: [10.1109/TSP.2016.2580524](https://doi.org/10.1109/TSP.2016.2580524).
- [11] GAO Wei, LI Jingchun, ZHOU Guangtao, *et al.* Adaptive Kalman filtering with recursive noise estimator for integrated SINS/DVL systems[J]. *The Journal of Navigation*, 2015, 68(1): 142–161. doi: [10.1017/S0373463314000484](https://doi.org/10.1017/S0373463314000484).
- [12] ZHU Hao, LEUNG H, and HE Zhongshi. State estimation in unknown non-Gaussian measurement noise using variational Bayesian technique[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2013, 49(4): 2601–2614. doi: [10.1109/TAES.2013.6621839](https://doi.org/10.1109/TAES.2013.6621839).
- [13] CHEN J and MA L. Particle filtering with correlated measurement and process noise at the same time[J]. *IET Radar, Sonar & Navigation*, 2011, 5(7): 726–730. doi: [10.1049/iet-rsn.2010.0365](https://doi.org/10.1049/iet-rsn.2010.0365).
- [14] SAHA S and GUSTAFSSON F. Particle filtering with dependent noise processes[J]. *IEEE Transactions on Signal Processing*, 2012, 60(9): 4497–4508. doi: [10.1109/TSP.2012.2202653](https://doi.org/10.1109/TSP.2012.2202653).
- [15] WANG Zongyuan and ZHOU Weidong. Robust linear filter with parameter estimation under student-t measurement distribution[J]. *Circuits, Systems, and Signal Processing*, 2019, 38(6): 2445–2470. doi: [10.1007/s00034-018-0972-8](https://doi.org/10.1007/s00034-018-0972-8).

王宗原: 男, 1977年生, 讲师, 研究方向为统计信号检测与处理。

周卫东: 男, 1966年生, 教授, 研究方向为卫星导航、组合导航技术。