一种基于二元估计与粒子滤波的故障预测算法

张 磊 李行善 于劲松 高占宝

(北京航空航天大学 自动化科学与电气工程学院, 北京 100191)

摘 要: 假设对象系统的故障演化过程可以由一个含有未知缓变参数的状态空间模型加以描述,则故障预测问题就可以转化为一个在已知当前系统信息的条件下,对系统未来某一时刻的状态变量的估计问题.针对该问题的求解提出了一种基于二元估计和粒子滤波的故障预测算法.算法的实施分为两个主要阶段:在状态估计阶段,采用两个并联的粒子滤波器迭代估计当前时刻对象系统故障演化模型状态和未知参数的后验分布.在状态预测阶段,对当前时刻故障演化模型状态的后验分布进行迭代采样,以采样样本粒子来近似估计未来时刻的状态变量的先验分布密度.在上述计算结果的基础上,结合相应的故障判据,算法采用计算对象系统未来时刻故障概率的方法预测其剩余使用寿命.仿真实验中将本文提出的算法与基于联合估计的故障预测算法进行对比,实验结果证明了所提算法的有效性.

关 键 词: 故障预测; 粒子滤波; 联合估计; 二元估计

中图分类号: TP 206^{+}.3

文献标识码: A **文 章 编 号**: 1001-5965(2008)07-0798-05

Fault prognostic algorithm based on dual estimation and particle filter

Zhang Lei Li Xingshan Yu Jinsong Gao Zhanbao

(School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100191, China)

Abstract: Supposing that system fault evolution process is expressed by a state space model with unknown slow time-variant parameters, then the fault prognostics can be formulated as a problem of estimating model states for some future time while knowing all the information about system till time step now. An algorithm based on dual estimation and particle filter was presented. This algorithm includes two major stages: on the state estimation stage, it used an iterative procedure to estimate the posterior distributions of system states and parameters alternatively based on two parallel connected particle filters; on the state prediction stage, the algorithm sampled former estimated posterior distributions iteratively and used the sampled particles to form the prior distributions of system states for some future time. Based upon above calculated results and combined with certain fault criterions, the time to failure could then be inferred by computing the probability of system failure. Comparing with the joint estimation approach, the simulation result demonstrates the validity and feasibility of proposed algorithm.

Key words: fault prognostics; particle filter; joint estimation; dual estimation

对于故障预测问题的研究,一种解决方案是采用基于对象故障演化模型状态估计方法.它的基本思想是:假设对象系统的故障演化过程可以由一个状态空间模型加以表述,即对象系统一些包含故障演化信息的状态变量(通常称为系统故障指征)与其故障演化过程之间隐含有某种对应关系,通过对系统状态变量未来一段时间内变化情况的预测,结合一定的判别准则,可以预知对象

收稿日期: 2007-06-20

基金项目: 总装"十一五"预研资助项目(51317050302)

作者简介: 张 磊(1981—),男,江西上饶人,博士生,zl666_81@126.com.

系统的故障情况和其剩余使用寿命.在该假设前提下,故障预测问题就转化为一个在已知当前系统信息,对系统未来某一时刻状态变量的估计问题.对于上述问题,贝叶斯估计方法从理论上给出了其求解方案,但对于大多数实际系统很难求得贝叶斯估计的解析解.因此,实际应用过程中出现了以卡尔曼滤波、扩展卡尔曼滤波和无迹滤波为代表的具体实现算法[1-3].然而,这些预测算法存在的主要的缺陷是系统状态变量必须满足高斯分布假设,这对于实际应用中广泛存在的非线性,非高斯分布情况的适用性存在很大限制.相对于以上方法,粒子滤波算法具有不受模型线性、高斯假设的约束的特点,理论上能适用于任意非线性非高斯动态系统的状态估计问题.因此,本文正是沿着采用粒子滤波算法实现状态估计问题的求解这一思路对故障预测问题作进一步深入地研究.

1　故障预测问题的数学描述

假设对象系统的故障演化过程可以采用如下的状态空间模型式(1)、式(2)[4]描述:

$$\boldsymbol{X}_k = f_{k-1}(\boldsymbol{X}_{k-1}, \boldsymbol{\theta}_k, \boldsymbol{U}_{k-1}, \boldsymbol{w}_{k-1}) \tag{1}$$

$$\boldsymbol{Y}_k = h_k(\boldsymbol{X}_k, \boldsymbol{\theta}_k, \boldsymbol{v}_k) \tag{2}$$

式(1)和式(2)分别为系统的状态方程和观测方程.其中,$\boldsymbol{X}_k$,$\boldsymbol{Y}_k$ 和 $\boldsymbol{U}_k$ 分别为系统的状态变量、观测变量和输入变量;$\boldsymbol{\theta}_k$ 为模型的未知参数(通常随时间缓慢变化),用于描述由于对象系统故障演化而导致的部分模型参数随时间变化特性;$\boldsymbol{w}_k$ 和 $\boldsymbol{v}_k$ 分别为系统噪声和观测噪声.对于上述的对象系统,故障预测问题可以转化为一个已知系统从某一起始时刻 0 至当前时刻 k 的观测序列$\{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \cdots, \boldsymbol{Y}_k\}$,同时在系统存在未知缓变参数的前提条件下,对未来某一时刻 $k+p(p>0)$的系统状态变量 $\boldsymbol{X}_{k+p}$求取其在一定的准则意义下的估计值的状态估计问题.

2　故障预测算法设计

针对上述的故障预测问题,采用粒子滤波算法实现其故障预测的过程相应地可以划分以下两个主要步骤:

1) 对当前时刻 k 的状态变量进行估计

由于对象系统模型中存在未知缓变参数,它的取值将对状态变量估计产生影响.对此问题的一种最直接的解决方案是采用所谓的联合估计方案,即将对象系统状态变量和未知参数作状态增广,在增广系统可观测的前提条件下,对增广系统做状态估计.该方案存在的一个问题是:增广系统状态将不再是各态遍历,其状态估计滤波器的一致收敛性不再得到保证,即可能无法同时跟踪原系统状态和未知参数的变化,从而直接影响预测算法的输出结果.对上述问题的一种解决方案是采用所谓的二元估计方案,即采用两个并联工作的滤波器依次迭代估计原系统的状态变量和未知参数的后验分布.二元估计算法可以视为期望最大算法(EM,Expectation Maximization)的一种特例,在 E 步估计原系统状态变量,在 M 步辨识原系统未知参数;E 步和 M 步的依次迭代执行,保证估计结果能逐渐收敛于状态和参数的真实值.二元估计算法中状态变量估计和模型参数辨识的对象模型分别为[4]

$$\left.\begin{aligned}\boldsymbol{X}_k &= f_{k-1}(\boldsymbol{X}_{k-1}, \boldsymbol{\theta}_k, \boldsymbol{U}_{k-1}, \boldsymbol{w}_{k-1})\\ \boldsymbol{Y}_k &= h_k(\boldsymbol{X}_k, \boldsymbol{\theta}_k, \boldsymbol{v}_k)\end{aligned}\right\} \tag{3}$$

$$\left.\begin{aligned}\boldsymbol{\theta}_k &= \boldsymbol{\theta}_{k-1} + \boldsymbol{\zeta}_{k-1}\\ \boldsymbol{Y}_k &= h_k(\boldsymbol{X}_k, \boldsymbol{\theta}_k, \boldsymbol{v}_k)\end{aligned}\right\} \tag{4}$$

在用式(3)估计系统状态时,参数 $\boldsymbol{\theta}_k$ 为前次模型参数估计的期望值;在用式(4)估计系统未知参数时,状态变量 $\boldsymbol{X}_k$ 为前次状态估计的期望值.在式(4)中,原对象系统时变参数 $\boldsymbol{\theta}_k$ 服从随机游走模型,模型参数 $\boldsymbol{\zeta}_k \sim N(0, \boldsymbol{\Sigma}_\zeta)$服从均值为 0,方差为 $\boldsymbol{\Sigma}_\zeta$ 的高斯随机分布.

由于未知参数 $\boldsymbol{\theta}_k$ 的估计对预测结果的准确性有直接的影响,同时由于参数本身的特性(未知参数可能在较长时间保持不变或只发生微小变化),使得所采用的粒子滤波算法不可避免地会面临粒子退化问题.为此,本文在重要性采样重采样(SIR,Sampling Importance Resampling)粒子滤波算法的再采样过程中加入马尔可夫链蒙特卡罗(MCMC,Markov Chain Monte Carlo)随机移动过程,以保持粒子的多样性,在一定程度上克服粒子贫化问题.并且,为了保证二元估计算法执行的效率,采用基本的 SIR 粒子滤波算法估计原对象系统的状态变量,以减少算法的运算量.

2)对未来时刻 $k+p$ 的状态变量进行预测

在采用粒子滤波方法获取的 k 时刻状态变量滤波密度 $p(\boldsymbol{X}_k|\boldsymbol{Y}_{1:k})$近似测度的基础上,可以估算出未来 $k+p$ 时刻的系统状态的先验概率密度 $p(\boldsymbol{X}_{k+p}|\boldsymbol{Y}_{1:k})$.参照文献[5],在一定的假设条件下,只需要采用 k 时刻对象系统的状态方程(系统未知参数 $\boldsymbol{\theta}_k$ 保持 k 时刻估计期望值不变)以滤波密度 $p(\boldsymbol{X}_k|\boldsymbol{Y}_{1:k})$的样本粒子为初始值,迭代采样出 $k+p$ 时刻系统状态变量的样本粒子 $\boldsymbol{X}_{k+p}^i$,同时保持其对应的权值不变.则 $k+p$ 时刻系统状

态变量的先验概率密度可以近似表述为

$$p(\boldsymbol{X}_{k+p} \mid \boldsymbol{Y}_{1:k}) \approx \sum_{i=1}^{N} W_k^i \delta(\boldsymbol{X}_{k+p} - \boldsymbol{X}_{k+p}^i) \quad (5)$$

$k+p$ 时刻状态变量的概率密度的预测结果 $p(\boldsymbol{X}_{k+p} \mid \boldsymbol{Y}_{1:k})$ 可以用来计算预测决策所需要的各项统计指标，如：故障概率、期望、方差等. 假设系统故障状态空间为 ϕ，则 $k+p$ 时刻故障概率为

$$\mathrm{Prob}(\boldsymbol{X}_{k+p}) = \mathrm{Prob}\{\boldsymbol{X}_{k+p} \in \phi\} = \int_{\boldsymbol{X}_{k+p} \in \phi} p(\boldsymbol{X}_{k+p} \mid \boldsymbol{Y}_{1:k}) \mathrm{d}\boldsymbol{X}_{k+p} = \sum_{i=1}^{S} W_k^i$$

其中 $S = \mathrm{number}\{\boldsymbol{X}_{k+p}^i \in \phi\}$.

$k+p$ 时刻状态变量的期望为

$$\mathrm{Mean}(\boldsymbol{X}_{k+p}) = E(\boldsymbol{X}_{k+p}) = \int_{\boldsymbol{X}_{k+p} \in \phi} (\boldsymbol{X}_{k+p} \mid \boldsymbol{Y}_{1:k}) \mathrm{d}\boldsymbol{X}_{k+p} = \sum_{i=1}^{N} \boldsymbol{X}_{k+p}^i W_k^i$$

$k+p$ 时刻状态变量的方差为

$$\mathrm{Var}(\boldsymbol{X}_{k+p}) = E\{[\boldsymbol{X}_{k+p} - E(\boldsymbol{X}_{k+p})][\boldsymbol{X}_{k+p} - E(x_{k+p})]^{\mathrm{T}}\} = \sum_{i=1}^{N} \{W_k^i [\boldsymbol{X}_{k+p}^i - \mathrm{Mean}(\boldsymbol{X}_{k+p})] \cdot [\boldsymbol{X}_{k+p}^i - \mathrm{Mean}(\boldsymbol{X}_{k+p})]^{\mathrm{T}}\}$$

3 故障预测仿真实验

采用文献[6]中的三罐系统作为故障预测仿真实验的对象，三罐系统的数学模型为

$$\left.\begin{aligned} \frac{\mathrm{d}\boldsymbol{X}}{\mathrm{d}t} &= \boldsymbol{A}(\boldsymbol{X}) + \boldsymbol{B}\boldsymbol{U}(t) \\ \boldsymbol{Y} &= [x_1 \quad x_2 \quad x_3]^{\mathrm{T}} \end{aligned}\right\} \quad (6)$$

其中 $\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \\ h_2 \end{bmatrix}$ $\boldsymbol{U} = \begin{bmatrix} Q_1 \\ Q_2 \end{bmatrix}$

$$\boldsymbol{A}(\boldsymbol{X}) = \frac{1}{\boldsymbol{A}} \begin{bmatrix} -Q_{13} \\ Q_{32} - Q_{20} \\ Q_{13} - Q_{32} \end{bmatrix} \qquad \boldsymbol{B} = \frac{1}{\boldsymbol{A}} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

上述模型中的未知系数 Q_{13}，Q_{32}，Q_{20} 分别为

$$Q_{13} = az_1 S_n \mathrm{sgn}(h_1 - h_3)(2g \mid h_1 - h_3 \mid)^{1/2}$$

$$Q_{32} = az_3 S_n \mathrm{sgn}(h_3 - h_2)(2g \mid h_3 - h_2 \mid)^{1/2}$$

$$Q_{20} = az_2 S_n (2gh_2)^{1/2}$$

三罐系统模型有关参数的正常取值如下：$A = 0.0154\ \mathrm{m}^2$，$S_n = 5 \times 10^{-5}\ \mathrm{m}^2$，$g = 9.81\ \mathrm{m/s}^2$，$az_1 = 0.5$，$az_2 = 0.6$，$az_3 = 0.5$. $\Delta t = 0.054\ \mathrm{s}$ 为仿真实验的时间间隔；整个仿真实验时间为 $200\Delta t$. 3个水槽的液位初始值分别为：$h_1^0 = 0.4\ \mathrm{m}$，$h_2^0 = 0.3\ \mathrm{m}$，$h_3^0 = 0.35\ \mathrm{m}$；系统的输入 Q_1，Q_2 分别为：$Q_k = 4.5 \times 10^{-5}\ \mathrm{m}^3/\mathrm{s}$，$k = 1, 2$. 系统的故障状态：

$$\phi = \left\{ \left| \frac{h_k - h_k^0}{h_k^0} \right| \geqslant 0.1 \right\} \qquad k = 1, 2, 3 \quad (7)$$

在 $100\Delta t$ 时，系统由于发生故障导致参数 az_2 按照以下的规律变化：

$$az_2 = \begin{cases} az_2^0 + (k - 100) \times 0.06 & 100 < k < 130 \\ 2.4 & k > 130 \end{cases} \quad (8)$$

假设实验中系统噪声服从非高斯分布 $w(k) \sim N(0, 1 \times 10^{-6}) + U(-10^{-3}, 10^{-3})$，量测噪声服从高斯分布 $v(k) \sim N(0, 1 \times 10^{-6})$. 采用联合估计和本文的二元估计算法进行故障预测，并且将其结果进行对比分析. 分别计算模型参数 az_2 的跟踪结果以及连续50步状态变量 h_2 预测的期望结果（分别显示时刻100～180之间间隔20个单位采样时间的计算结果），分别如图1～图3.

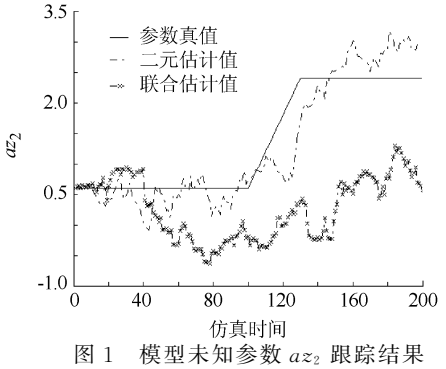


图1 模型未知参数 az_2 跟踪结果

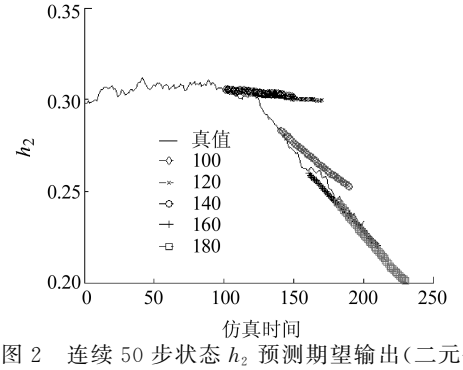


图2 连续50步状态 h_2 预测期望输出（二元估计）

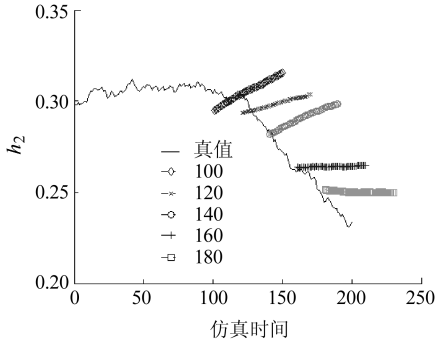


图3 连续50步状态 h_2 预测期望输出（联合估计）

对比上述结果可以看出，在该仿真实验中采用基于联合估计的故障预测算法基本失效了，这

一点可以从其对参数的跟踪结果以及其对状态变量发展趋势的预测结果中清晰的反映出来(算法对参数变化跟踪结果发散,对状态发展趋势预测与实际情况有很大的偏差).分析其可能的原因是由于原对象状态变量和未知参数组成的增广系统对未知参数不可观测,即算法的应用前提得不到满足造成.

与此同时,仿真结果证明采用本文的基于二元估计的故障预测算法能较好的跟踪模型未知参数的变化,同时其对状态发展趋势预测结果与实际情况吻合的很好.

采用本文的故障预测算法计算连续 10 步的故障概率预测输出,计算部分结果如表 1 所示(k 为当前时刻,p 为预测步长).

在时刻 145 之前的时间段内,预测的故障概率很小.该结果包含两种情况:

1) 时刻 100 之前,系统工作正常;

2) 时刻 101～145 之间,系统虽然发生故障,但其故障趋势不是很明显(参数 az_2 逐渐偏离其正常值),同时故障程度积累到一定的量值需要一定时间,因此在该时段,故障预测结果为系统工作正常.

从时刻 146 至系统故障时刻 161,故障概率预测值发生急剧变化.在此对其进行一定的处理,从而更加清晰的反映系统故障的变化情况,以便进行故障预测的决策.一种具体的处理方法如下(计算加权故障概率):

$$\widehat{\mathrm{Prob}}(\boldsymbol{X}_{k+p})_t=\sum_{i=t}^{N}\mathrm{Prob}(\boldsymbol{X}_{k+p-i},i)\times W_i$$

其中,$1\leqslant t<N$,N 表示对 $k+p$ 时刻预测决策时发生影响的最大步长(实际应用中通常是由经验数据决定);$\widehat{\mathrm{Prob}}(\boldsymbol{X}_{k+p})_t$ 表示 $k+p$ 时刻预测步长为 t(预测决策的实际步长)的加权故障概率预测结果;$\mathrm{Prob}(\boldsymbol{X}_{k+p-i},i)$ 表示当前时刻为 $k+p-i$,预测步长为 i 的故障概率;W_i 为其相应的权值.数据处理结果如表 2 所示.

表 1　故障概率预测部分结果(1～10 步预测输出)

k	p									
	1	2	3	4	5	6	7	8	9	10
146	0.07	0.19	0.35	0.55	0.61	0.73	0.81	0.88	0.88	0.89
147	0.06	0.26	0.39	0.51	0.65	0.67	0.75	0.76	0.80	0.81
148	0.27	0.5	0.56	0.74	0.77	0.85	0.92	0.94	0.94	0.94
149	0.37	0.53	0.7	0.75	0.81	0.88	0.92	0.92	0.96	0.95
150	0.50	0.78	0.71	0.76	0.79	0.80	0.84	0.85	0.85	0.87
151	0.63	0.75	0.86	0.89	0.87	0.92	0.96	0.98	0.97	0.96
152	0.65	0.71	0.73	0.80	0.93	0.94	0.95	0.97	0.98	0.98
153	0.51	0.64	0.74	0.74	0.84	0.83	0.87	0.92	0.92	0.92
154	0.51	0.63	0.71	0.81	0.85	0.85	0.91	0.92	0.93	0.94
155	0.66	0.75	0.68	0.73	0.80	0.84	0.83	0.88	0.94	0.95
156	0.44	0.51	0.52	0.64	0.74	0.77	0.83	0.87	0.88	0.87
157	0.19	0.34	0.43	0.57	0.59	0.7	0.68	0.72	0.76	0.81
158	0.12	0.16	0.37	0.39	0.5	0.56	0.61	0.68	0.7	0.7
159	0.41	0.6	0.65	0.74	0.76	0.87	0.8	0.82	0.87	0.83
160	0.63	0.7	0.72	0.8	0.86	0.87	0.86	0.86	0.91	0.91
161	0.97	0.9	0.98	0.96	0.97	0.93	0.93	0.93	0.96	0.97

表 2　加权故障概率预测结果

p	$k+p$					
	156	157	158	159	160	161
10	0.89	0.81	0.94	0.95	0.87	0.96
9	0.84	0.87	0.95	0.90	0.94	0.97
8	0.87	0.89	0.91	0.92	0.93	0.95
7	0.88	0.87	0.92	0.93	0.92	0.94
6	0.87	0.88	0.93	0.91	0.90	0.92
5	0.87	0.89	0.91	0.90	0.90	0.89
4	0.86	0.87	0.90	0.87	0.85	0.84
3	0.84	0.85	0.87	0.83	0.80	0.78
2	0.82	0.84	0.83	0.77	0.72	0.76
1	0.80	0.80	0.76	0.71	0.69	0.75

上述处理的物理意义如下:用于 $k+p$ 时刻预测决策(提前 t 个时刻的预测)的故障概率值(加权故障概率)是由其之前的一段时间之内($N-t+1$ 个时刻),对其后的 $k+p$ 时刻故障概率预测结果加权决定,其权值反映相应时刻故障概率预测结果对最终决策的影响程度.

若将系统故障决策判据设定为加权故障概率预测大于 0.9,则可以预知故障的发生时刻大约在时刻 158～161 之间,与实际值 161 很接近.同时,从表 1 和表 2 中还可以看出,时刻 156～160 故障概率预测结果与实际情况之间的偏差逐渐增

加(反映在表2的后4组数据中更加明显),分析其原因主要是由于在该段时间内,算法对模型未知参数跟踪结果出现了一个较大的振荡,导致跟踪结果偏离了当前参数的真值(为了有效跟踪仿真试验中模型参数变化,参数随机游走模型中噪声方差设置较大),从而直接影响了算法预测结果的准确性.

在时刻161之后,模型故障趋势十分明显,故障预测结果几乎全部为1,因此可以断定故障已经发生.

以上仅演示了一个简单的故障预测决策过程,目的是为了说明本文所提算法的有效性.在实际的故障预测应用中,其预测决策过程一定是一个对多种有用信息进行有效融合的过程,决策结果必定是通过对计算的多种故障特征(如故障概率、期望等)进行综合评价后给出的.因此,在预测决策过程中如何有效的应用信息融合技术,将成为故障预测一个十分值得研究的内容.

4 结 论

仿真计算结果证明本文提出的基于二元估计和粒子滤波的故障预测算法能很好地满足对象系统非线性、非高斯噪声状态下故障预测问题的求解需求.将本文提出的方法与基于联合估计的预测方案相结合,能对一类实际的故障预测问题进行有效的求解.但该方法也存在不足,采用两个滤波器交替工作的机制导致基于二元估计的故障预测算法的运算量大大高于基于联合估计的预测方法;对模型未知参数的跟踪结果直接影响预测结果,如何有效的跟踪模型参数,同时减小其跟踪过程中的振荡,是算法进一步改进时的一个重要的研究方向.此外,故障预测的决策过程中有效的应用信息融合技术也是故障预测的一个重要研究内容.

参考文献(References)

[1] Ray A, Tangirala S. Stochastic modeling of fatigue crack dynamics for on-line failure prognostics[J]. IEEE Transactions on Control Systems Technology, 1996, 4(4):443-451

[2] Swanson D C. A general prognostic tracking algorithm for predictive maintenance[C]//Proceedings of IEEE Aerospace Conference. Big Sky, Mantana: IEEE, 2001:2971-2977

[3] Chelidze D. Multimode damage tracking and failure prognosis in electromechanical system[C]//Willett P K, Kirubarajan T. Components and Systems Diagnostics, Prognostics, and Health Management II, Proceedings of SPIE. Orlando: SPIE, 2002: 1-12

[4] Wan E A, Van Der Merwe R. The unscented Kalman filter for nonlinear estimation[C]//Adaptive Systems for Signal Processing, Communication, and Control Symposium 2000. Lake Louise, Alta: IEEE, 2000:153-158

[5] Doucet A, Godsill S J, Andrieu C. On sequential Monte Carlo sampling methods for Bayesian filtering[J]. Statistics and Computing, 2000, 10(3):197-208

[6] Xie X Q, Zhou D H, Jin Y H. Strong tracking filter based adaptive generic model control[J]. Journal of Process Control, 1999, 9(4):337-350

(责任编辑:赵海容)