

非线性系统中目标跟踪性能 评估的新度量

毛艳慧^{1,2} 韩崇昭^{1,2}

摘要 在将误差谱度量推广应用到动态系统滤波器的评估时, 整个时间轴上将绘出一个三维图, 不直观且不利于分析. 为了解决这一难题, 提出一个新的度量——动态误差谱, 并给出了三种不同应用背景下的计算形式. 其中一种采用了几何平均的形式因而可以给出较为“公正”的评估结果, 可直接应用于动态系统. 接着从新的角度把这一度量应用于对四种非线性滤波算法的评估, 仿真结果验证了动态误差谱这一新度量的有用性和有效性.

关键词 非线性量测, 目标跟踪, 性能评估, 动态误差谱, 几何平均

引用格式 毛艳慧, 韩崇昭. 非线性系统中目标跟踪性能评估的新度量. 自动化学报, 2014, 40(11): 2650–2653

DOI 10.3724/SP.J.1004.2014.02650

A New Metric for Target Tracking Performance Evaluation of Nonlinear Systems

MAO Yan-Hui^{1,2} HAN Chong-Zhao^{1,2}

Abstract When applying error spectrum measure to performance evaluation for filtering of dynamic systems, it will plot a 3D figure over the total time span, which is not intuitive and difficult to be analyzed. In this study, to overcome its drawbacks, a new metric, dynamic error spectrum, is proposed to summarize the ES curve. Three forms under different application backgrounds are given, one of which is evenly taking into account both good and bad behaviors of an estimator and so can provide more impartial evaluation results. It can be applied to a variety of dynamic systems directly. Then from new perspectives, four nonlinear filters are chosen to illustrate the superiority of the metric. Simulation results validate its utility and effectiveness.

Key words Nonlinear measurements, target tracking, performance evaluation, dynamic error spectrum, geometric mean

Citation Mao Yan-Hui, Han Chong-Zhao. A new metric for target tracking performance evaluation of nonlinear systems. *Acta Automatica Sinica*, 2014, 40(11): 2650–2653

现有对非线性量测滤波算法进行性能比较及评估的研究中^[1], 一个广泛应用的指标是均方根误差 (Root mean square error, RMSE). 这一指标很容易受大的误差值主导. 为克服这一缺点, 文献 [2] 提出了一批新的可供选择的绝对度量, 如

调和平均误差 (Harmonic average error, HAE)、几何平均误差 (Geometric average error, GAE)、平均欧几里得误差 (Average Euclidean error, AEE)、误差中位数和误差的众数. 但因为这些准则都只侧重反映估计性能的某一个方面, 所以它们是片面的. 甚至在比较两个估计器性能时, 会给出相悖的比较结果. 也就是说, 给定两组估计器性能 A 组和 B 组, 用一个度量评估时 A 组好; 可换另一个度量时, 可能评估结果反而 B 组好. 因此, 就希望有一个综合全面的指标. 文献 [3] 综合了上面的指标, 提出了一个综合性的度量——误差谱. 误差谱可以提供一条估计性能的曲线, 这条曲线包括了调和平均、几何平均、算数平均和广泛应用的均方根误差, 所以这一度量可以同时说明不同方面的估计性能.

但是, 误差谱是直接适用于参数估计的形式给出的. 对动态系统, 在任意时刻这一度量将给出状态估计器的性能曲线, 因此, 在整个时间轴上, 误差谱将绘出一个三维图. 尽管可以用计算机借助一些绘图工具绘出这一三维图, 但非常不直观, 尤其是当两个估计器的误差谱曲线在某些时刻相交而又需要知道哪个估计器有更好的性能时. 因此, 怎样将多条曲线组合得到一个二维图, 使误差谱可以推广到动态系统的问题, 亟待解决.

本文给出了一个动态误差谱的度量, 用三种形式来解决上面提到的问题. 如果给每个时刻误差谱曲线的各个点赋不同的权值, 加权平均是一个可用的方法, 每一个时刻的误差谱曲线将变成一个点, 在整个时间轴上可以绘出一条线. 但在实际中, 很难合理地设计这些权值; 本文给出的另一个方法是通过计算误差谱曲线下的面积来概括误差谱曲线. 但这一方法是有缺陷的, 因为它本质上是误差谱曲线上点的算数平均, 所以很有可能误差谱曲线上大的值主导; 最后, 本文给出了一个更合理的几何平均形式的方法. 这个方法充分利用了几何平均的平衡性质, 并且可以给出更公平合理的评估结果. 即任何大的值都能被一个足够小的值平衡, 反之亦然. 由于这一特性, 这一方法可以很好地平衡估计器的好坏性质, 因此最终的性能评估结果也更有说服力, 而且可以直接应用于动态系统.

目前, 已有大量针对非线性滤波的研究成果^[4–7]. 本文针对非线性量测系统, 选择了经典的无偏量测转换 (Unbiased converted measurement, UCM)^[8]、最优线性无偏估计 (Based linear unbiased estimator, BLUE)^[9]、扩展卡尔曼滤波器 (Extended Kalman filter, EKF) 和无迹滤波器 (Unscented filter, UF)^[10–11] 在不同场景下进行性能上的分析和比较, 说明动态误差谱这一度量优于传统的均方根误差, 仿真结果很好地阐明了本文提出的动态误差谱的有效性.

1 动态误差谱

1.1 误差谱

\mathbf{x} 、 $\hat{\mathbf{x}}$ 分别代表待估量和估计值. 给定一个估计器在 k 时刻的估计误差 $\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$, $e = \|\tilde{\mathbf{x}}\|$ 或 $e = \|\tilde{\mathbf{x}}\|/\|\mathbf{x}\|$ 代表绝对的或相对的估计误差, $\|\cdot\|$ 可以是 1-范数或 2-范数. 以下标 k 和 r 分别代表时刻值和误差谱曲线的自变量. 定义误差谱 (Error spectrum, ES)^[3]

$$ES_{k,r} = S_k(r) = [E(e_k^r)]^{\frac{1}{r}} = \left(\int e_k^r dF(e_k) \right)^{\frac{1}{r}} =$$

收稿日期 2013-08-22 录用日期 2014-04-15
Manuscript received August 22, 2013; accepted April 15, 2014
国家自然科学基金创新研究群体科学基金 (61221063), 国家自然科学基金 (61074176)

Supported by Foundation for Innovative Research Groups of National Natural Science Foundation of China (61221063) and National Natural Science Foundation of China (61074176)

本文责任编辑 丛爽
Recommended by Associate Editor CONG Shuang

1. 西安交通大学智能网络与网络安全教育部重点实验室 西安 710049 2. 西安交通大学电信学院综合自动化研究所 西安 710049

1. Ministry of Education Key Laboratory of Intelligent Networks and Network Security (MOE KLINNS Lab), Xi'an Jiaotong University, Xi'an 710049 2. Institute of Integrated Automation, School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049

$$\begin{cases} (\int e_k^r f(e_k) de_k)^{\frac{1}{r}}, & \text{连续 } e \\ (\sum e_{k,i}^r p_i)^{\frac{1}{r}}, & \text{离散 } e \end{cases} \quad (1)$$

式中, $r \in \mathbf{R}^+ = [-\infty, \infty]$ ^[3], $F(e_k)$ 、 $f(e_k)$ 和 p_i 分别是误差 e_k 的累积分布函数、概率密度函数和分布律. 由于 r 的取值范围为 $-\infty \sim +\infty$, 所以 $S(r)$ 可以绘出整个 r 轴的二维图, 并且可以解释估计器取所有不同 r 值的估计性能. 显然, 误差谱曲线越低, 相应的估计误差越小, 估计器的性能越好. 另外, 区间 $r \in [-1, 2]$ 涵盖了文献 [2] 提到的性能评估准则:

- $S(-1) = \text{HAE}$
- $S(0) = \text{GAE}$
- $S(1) = \text{AEE}$
- $S(2) = \text{RMSE}$

传统的均方根误差度量仅是误差谱曲线上在 $r = 2$ 时的一个点^[3].

误差谱度量可以综合考虑大和小估计误差, 而传统的均方误差度量却只关注大的估计误差, 所以很容易受大误差主导. 在期望对滤波器的性能有一个公正综合的评估时, 即评估结果既不受小误差值主导, 又不受大误差值主导, 须用误差谱这一综合性的度量. 但文献 [3] 是针对参数估计提出的这一度量. 对动态系统的状态估计器, 误差谱在任意时刻都是一条曲线, 在整个时间轴上绘出的将是一个三维图, 不直观且存在分析上的困难. 尤其当两个估计器的误差谱曲线在某些时刻有交点, 如图 1 所示, 就很难断定哪一个估计器的性能更好. 因此, 若要把这一度量扩展应用到动态系统, 亟须解决如何将多条曲线组合得到一个二维图, 期望得到随着时间变化, 动态系统的滤波器的综合性能. 由此本文引入动态误差谱的概念.

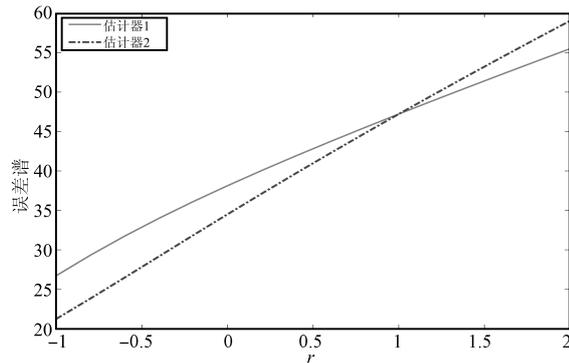


图 1 两个滤波器在 k 时刻相交的误差谱曲线

Fig. 1 The intersection of ES curves for two filters at time k

1.2 动态误差谱

记集合 $\{r_i\}_{i=1}^n$ 为评估时关心的 r 的集合, 给定 $\{r_i\}_{i=1}^n$, 如果有先验信息设计相应于每一个 r_i 的权值 $\{w_i\}_{i=1}^n$, 这里 $\sum w_i = 1$, 加权形式的动态误差谱 (Dynamic error spectrum, DES) 可以简单定义为

$$\text{DES}_k^W = \sum_{i=1}^n S_k(r_i) w_i \quad (2)$$

但是如果没有任何先验信息, 很难合理地设计这些权值, 解

决的方法是利用误差谱曲线下的面积, 定义如下:

$$\text{DES}_k^{\text{AM}} = \frac{1}{r_n - r_1} \int_{r_1}^{r_n} S_k(r) dr \approx \frac{1}{n} \sum_{i=1}^n S_k(r_i) \quad (3)$$

显然, 误差谱曲线越高, 曲线下的面积越大, 相应的估计器性能也就越差. 使用这一方法, 如图 1 所示, 在区间 $r \in [-1, 2]$ 上, 由于点划线下的面积明显小于实线下的面积, 估计器 2 的性能优于估计器 1 的性能.

尽管式 (3) 似乎解决了当两个估计器的误差谱曲线有交点时性能比较的问题, 但式 (3) 本质上是 $S_k(r_i)$ 的代数平均 (Arithmetic mean, AM), 因此仍可能受大 $S_k(r_i)$ 值的主导, 见图 2. 图 2 是 k 时刻两个估计器的误差谱曲线. 从图 2 可以看出, 虚线一直低于实线, 仅在很少的一段 r 值上高于实线, 但若采用曲线下的面积进行比较, 估计器 1 的性能反而高于估计器 2, 这是由于代数平均受大值主导造成的.

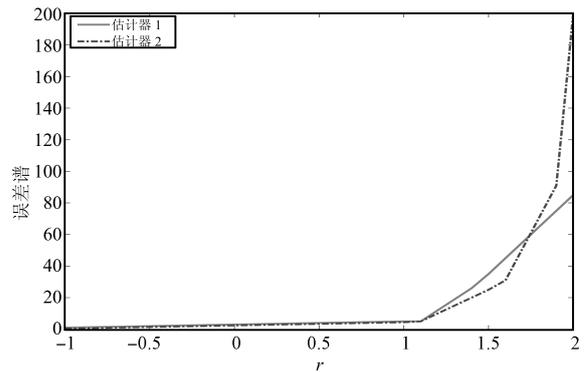


图 2 大误差主导下的误差谱曲线下的面积

Fig. 2 The domination of using the area under ES curve

几何均值 (Geometric mean, GM) 因为其既不受大误差主导也不受小误差主导, 所以是“平衡”的^[2]. 在对滤波器的性能综合评估时, 一个期望的度量应该同时考虑大和小估计误差, 即大误差和小误差能同时计算在内. 因此, 为了消除大值主导现象, 本文提出采用几何均值来概括每条误差谱曲线, 定义为:

$$\begin{aligned} \text{DES}_k^{\text{GM}} &= \exp \left(\frac{1}{r_n - r_1} \int_{r_1}^{r_n} \ln S_k(r) dr \right) \approx \\ &\exp \left(\frac{1}{n} \sum_{i=1}^n \ln S_k(r_i) \right) \end{aligned} \quad (4)$$

这一方法由于几何均值的“平衡”性质, 可以给出更“公正”的评估结果. 即与均方根误差和代数均值不同, 任何大的值都可以被一个足够小的值平衡, 反之亦然. 正是因为这一性质, 滤波器好和差估计结果都能被考虑和计算到, 评估结果也因此更有说服力. 同时, 它可以直接应用于动态系统.

式 (2)~(4) 给出的动态误差谱的三种形式可以直接应用到动态系统. 若只关注相应误差谱曲线上的某些点的滤波器的性能, 且相应的权值已知, 推荐使用加权形式; 若没有权值的先验信息, 且又期望得到滤波器的综合性能, 则可应用后两种形式. 一旦误差谱曲线在某些点上急剧爬升, 即大值主导现象存在, 则优先推荐应用几何均值形式. 第 2 节将详细说明这一度量的有效性.

2 仿真实例

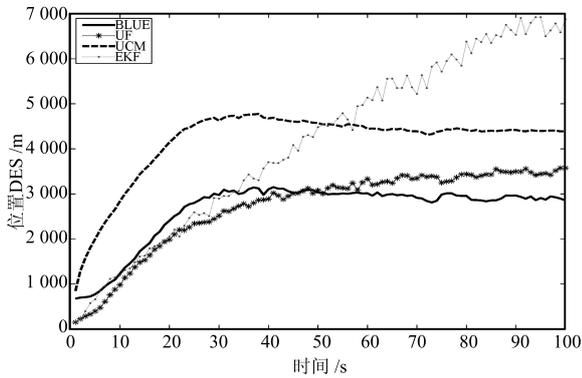
考虑二维的笛卡尔 $x-y$ 空间, 传感器位置在原点. 目标的采样间隔为 1 秒. 初始时刻目标的位置来自高斯分布, 均值为 [50 km, 200 km], 方向标准差为 5 km. 目标运动为近似匀速, 来自均值为 [1 km/s, 0 km/s], 各方向标准差为 0.1 km/s 的高斯分布.

误差谱的 r 取值区间为 $[-1, 2]$, 这是因为它涵盖了几个常用的度量指标. 本文将结合常用的均方根误差和 (Non credibility index, NCI) 指标^[12-13] 阐释动态误差谱度量在评估非线性滤波器性能的有效性. 本部分仿真采用 1000 次蒙特卡洛试验, 时间为 100 秒.

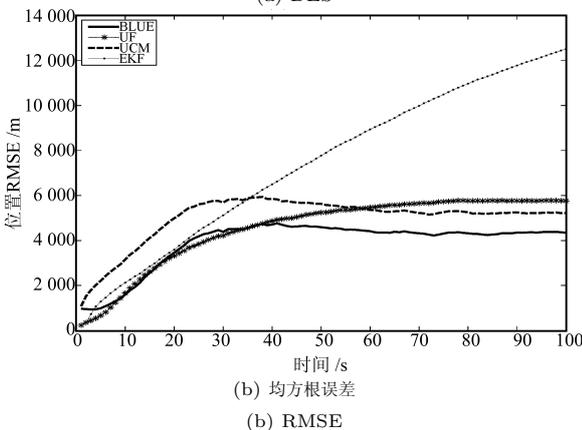
2.1 仿真结果与讨论

场景 1. 选取测量误差 $\sigma_r = 4\text{ m}$, $\sigma_\beta = 0.1\text{ rad}$.

从图 3(a) 可以看出, EKF 滤波器性能出现发散, 其 DES 曲线几乎线性上升, 另外三个滤波器受初值的影响, 在 40 秒之后开始收敛, 其中 BLUE 的性能最好, UF 次之, UCM 的动态误差谱曲线远高于 BLUE 和 UF, 这与图 4 所示 NCI 结果一致. 但是在图 3(b) 所示中, UCM 的 RMSE 曲线在大约 60 秒后低于 UF, 这与图 3(a) 是矛盾的, 且四个滤波器的 RMSE 曲线明显高于 DES.



(a) 动态误差谱
(a) DES



(b) 均方根误差
(b) RMSE

图 3 场景 1 下 4 种滤波器的动态误差谱和均方根误差比较
Fig. 3 Comparison with DES and RMSE of the four filters for Case 1

场景 2. 选取测量误差 $\sigma_r = 40\text{ m}$, $\sigma_\beta = 0.1\text{ rad}$.

在这一场景中, 增大距离误差标准差, 如图 5(a) 所示, UCM 性能最差. 与场景 1 中相比, EKF 滤波器性能有所改善, 有收敛的倾向, 这是因为径向距离的误差变大, 相比起

场景 1, 系统的非线性反而变弱. BLUE 滤波器的性能仍是最好. 从在图 5(b) 可以看出, EKF 也有收敛, 但性能最差, UCM 反而最后与 UF 的 RMSE 曲线重合, 这是因为 DES 考虑了 $r \in [-1, 2]$ 整个区间上的性能, 而 RMSE 只考虑了 $r = 2$ 的一个点.

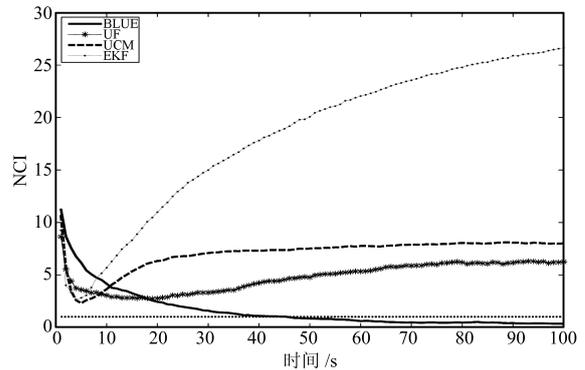
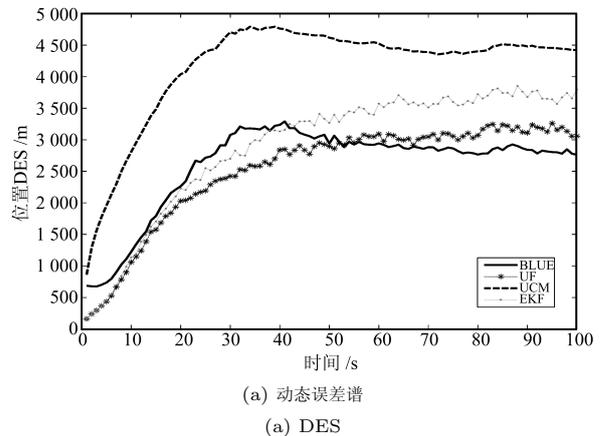
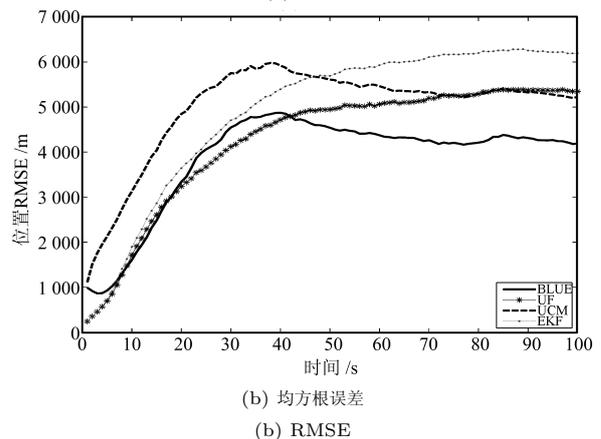


图 4 场景 1 下 4 种滤波器的 NCI 比较

Fig. 4 NCI comparison of the four filters for Case 1



(a) 动态误差谱
(a) DES



(b) 均方根误差
(b) RMSE

图 5 场景 2 下 4 种滤波器的动态误差谱和均方根误差比较
Fig. 5 Comparison with DES and RMSE of the four filters for Case 2

从图 6 可以看出, UCM 的性能最差, 与 DES 评估的结果一致.

评估结果显示, 本文提出的动态误差谱度量能很好地显示不同滤波器在不同场景下的整体性能, 且与 NCI 评估结果一致, 说明这一度量比均方根误差度量更综合也更有效.

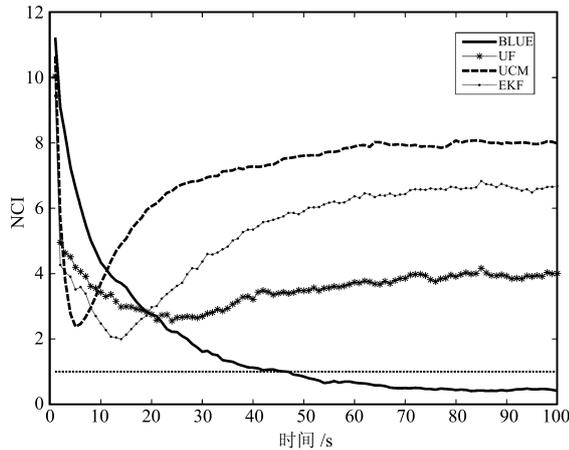


图6 场景2下4种滤波器的NCI比较

Fig.6 NCI comparison of the four filters for Case 2

3 结论

误差谱度量综合了很多绝对的误差指标, 在估计性能评估中绘出一条曲线, 因此更综合也更全面, 但是当把这一度量应用到动态系统时, 评估结果只能是三维图, 很不直观且不好分析. 为解决这一问题, 本文提出了一个新的指标——动态误差谱度量, 并给出了它的三种形式来计算每一个时刻的误差谱曲线. 其中一个方法利用了几何平均的形式, 因而是“平衡”的, 可以给出“公正”的评估结果. 这一度量可以直接应用于动态系统.

本文对EKF、UCM、UF、BLUE四种滤波器在两种场景下分别进行了性能分析和比较, 仿真结果验证了本文提出这一度量几何平均形式下的有效性和有用性.

References

- Xie Lei, Feng Hao, Zhang Jian-Ming. A new approach to performance assessment based on initial closed-system. *Acta Automatica Sinica*, 2013, **39**(5): 649–653
(谢磊, 冯皓, 张健明. 一种基于初始闭环系统的性能评估方法. *自动化学报*, 2013, **39**(5): 649–653)
- Li X R, Zhao Z L. Evaluation of estimation algorithms. part I: incomprehensive measures of performance. *IEEE Transactions on Aerospace and Electronic Systems*, 2006, **42**(4): 1340–1358
- Li X R, Zhao Z L, Duan Z S. Error spectrum and desirability level for estimation performance evaluation. In: *Proceedings of Workshop on Estimation, Tracking and Fusion: A Tribute to Fred Daum*. Monterey, CA, USA, 2007
- Wang Lu, Li Guang-Chun, Qiao Xiang-Wei, Wang Zhao-Long, Ma Tao. An adaptive UKF algorithm based on the maximum likelihood principle and expectation maximization algorithm. *Acta Automatica Sinica*, 2012, **38**(7): 1200–1210
(王璐, 李光春, 乔相伟, 王兆龙, 马涛. 基于极大似然准则和最大期望算法的自适应UKF算法. *自动化学报*, 2012, **38**(7): 1200–1210)
- Zhou Bo, Qian Kun, Ma Xu-Dong, Dai Xian-Zhong. A new nonlinear set membership filter based on guaranteed bounding ellipsoid algorithm. *Acta Automatica Sinica*, 2013, **39**(2): 150–158
(周波, 钱堃, 马旭东, 戴先中. 一种新的基于保证定界椭球算法的非线性集员滤波器. *自动化学报*, 2013, **39**(2): 150–158)
- Song Yu, Li Qing-Ling, Kang Yi-Fei, Yan De-Li. SLAM with square-root cubature Rao-Blackwillised particle filter. *Acta Automatica Sinica*, 2014, **40**(2): 357–367
(宋宇, 李庆玲, 康轶非, 闫德立. 平方根容积Rao-Blackwillised粒子滤波SLAM算法. *自动化学报*, 2014, **40**(2): 357–367)

- Lian Feng, Han Chong-Zhao, Li Chen. Multiple model GM-CBMeMber filter and track continuity. *Acta Automatica Sinica*, 2014, **40**(2): 336–47
(连峰, 韩崇昭, 李晨. 多模型GM-CBMeMber滤波器及航迹形成. *自动化学报*, 2014, **40**(2): 336–347)
- Mo L B, Song X Q, Zhou Y Y, Sun Z K, Bar-Shalom Y. Unbiased converted measurements for tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 1998, **34**(3): 1023–1027
- Zhao Z L, Li X R, Jilkov V P. Best linear unbiased filtering with nonlinear measurements for target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 2004, **40**(4): 1324–1336
- Julier S J, Uhlmann J K. New extension of the Kalman filter to nonlinear systems. In: *Proceedings of Aero Sense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*. Orlando, Florida: SPIE, 1997. 182–193
- Julier S J, Uhlmann J K. Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 2004, **92**(3): 401–422
- Li X R, Zhao Z L. Measuring estimator's credibility: non-credibility index. In: *Proceedings of the 2006 International Conference on Information Fusion*. Florence: IEEE, 2006. 1–8
- Li X R, Zhao Z L, Li X B. Evaluation of estimation algorithms: credibility tests. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 2012, **42**(1): 147–163

毛艳慧 西安交通大学电子与信息工程学院综合自动化研究所博士研究生. 主要研究方向为性能评估, 多源信息融合. 本文通信作者.

E-mail: yhmaoxjtu@gmail.com

(MAO Yan-Hui Ph.D. candidate at the Institute of Integrated Automation, School of Electronic and Information Engineering, Xi'an Jiaotong University. Her research interest covers performance evaluation and target tracking. Corresponding author of this paper.)

韩崇昭 西安交通大学电子与信息工程学院教授. 主要研究方向为多源信息融合, 随机控制与自适应控制, 非线性频谱分析.

E-mail: czhan@mail.xjtu.edu.cn

(HAN Chong-Zhao Professor at the School of Electronic and Information Engineering, Xi'an Jiaotong University. His research interest covers multi-source information fusion, stochastic control and adaptive control, and nonlinear spectral analysis.)