

文章编号: 1001-0920(2011)09-1402-05

## 马尔可夫链蒙特卡洛重要度采样与多目标跟踪

龙云利, 徐 晖, 安 玮

(国防科技大学 电子科学与工程学院, 长沙 410073)

**摘 要:** 针对强杂波环境下的多目标跟踪问题, 提出一种基于马尔可夫链蒙特卡洛重要度采样的跟踪方法. 通过马尔可夫链蒙特卡洛实现对联合关联事件的采样, 据此计算目标可关联量测数据的边缘关联概率. 在联合关联事件求解中利用单目标量测的概率密度进行重要度采样, 提高采样效率. 马尔可夫链蒙特卡洛重要度采样方法克服了联合概率数据关联中的“组合爆炸”问题, 能够在强杂波干扰下较好地实现多目标实时跟踪. 通过仿真实验对比分析了算法的跟踪精度和处理的时效性, 验证了方法的有效性.

**关键词:** 传感器; 多目标跟踪; 马尔可夫链蒙特卡洛; 重要度采样; 联合概率数据关联

**中图分类号:** TP391

**文献标识码:** A

## Markov chain Monte Carlo and importance sampling for multiple targets tracking

LONG Yun-li, XU Hui, AN Wei

(School of Electronic Science and Engineering, National University of Defense Technology, Changsha 410073, China. Correspondent: LONG Yun-li, E-mail: feiyunlyi@126.com)

**Abstract:** This paper presents an algorithm based on Markov chain Monte Carlo and importance sampling(MCMCIS) for tracking multi-target in a dense environment. The joint associated events are sampled by the Markov chain Monte Carlo and the marginal association probability of the measurement to the target is calculated. The probabilistic density is utilized when sampling the associated events as to improve the efficiency. Although the joint probabilistic data association(JPDA) is NP-hard, the MCMCIS provides the ability to track multi-target timely in a dense environment. The simulation experiments are implemented to analyze the tracking precision and processing time, which shows the effectiveness of the algorithm.

**Key words:** sensors; multi-target tracking; Markov chain Monte Carlo; importance sample; joint probabilistic data association

### 1 引 言

多目标跟踪广泛存在于军事监视、交通管制等众多光学传感器应用领域中. 多目标跟踪的本质是对含有杂波噪声干扰的量测信息进行处理, 最终实现对目标轨迹状态的估计, 其核心是数据关联, 解决量测信息和目标之间的模糊关系<sup>[1]</sup>. 按照跟踪过程中的最优化目标函数可将其分为 3 类<sup>[2]</sup>: 1) 启发式方法. 没有具体的优化目标函数, 如典型的最近邻滤波方法(NNF), 选择与预测目标位置最近的量测信息作为目标轨迹的当前观测进行滤波处理. 该方法只适合于低密度杂波干扰环境. 2) 极大后验概率方法(MAP). 以极大后验概率作为最优化目标函数, 确定量测与目标的关联最终实现目标轨迹的状态估计. 著名的多假

设跟踪(MHT)<sup>[3]</sup>及其改进算法<sup>[4]</sup>是一种多帧观测数据批处理的极大后验概率方法, 由于存在关联和目标状态的多维求解过程, 计算效率较低. 3) 贝叶斯方法. 精确的贝叶斯方法与 MAP 方法一样难以进行有效求解. 经典的联合概率数据关联<sup>[5]</sup>(JPDA) 方法是一种伪贝叶斯最优方法, 它根据目标和量测的可关联情况枚举联合关联事件并计算其概率, 对关联事件的滤波结果按照关联概率进行加权处理. JPDA 方法在低密度杂波环境下能够较好地实现多目标跟踪, 但在高密度杂波干扰下, 联合关联事件的计算非常复杂, 往往导致组合爆炸.

为解决 JPDA 的组合爆炸问题, Oh<sup>[2]</sup>等人提出了基于马尔可夫链蒙特卡洛(MCMC)的多目标单帧处

收稿日期: 2010-06-01; 修回日期: 2010-07-20.

基金项目: 国防装备预研基金项目(9140A21041110KG0148).

作者简介: 龙云利(1981—), 男, 博士生, 从事目标检测与跟踪的研究; 徐晖(1963—), 男, 教授, 博士生导师, 从事空间信息对抗等研究.

理策略, 并且理论分析和实验证明了MCMC相比MHT等算法在跟踪性能和处理时效上的优势, 但该方法在联合关联时间求解过程中对所有可关联量测数据进行均匀采样, 导致算法的收敛速度较低。

本文在MCMC的基础上进行研究, 提出了一种基于马尔可夫链蒙特卡洛与重要度采样的多目标跟踪方法。通过马尔可夫链蒙特卡洛实现对目标与量测关联事件的采样, 同时利用目标量测的概率密度进行重要度采样, 对关联事件进行统计获取量测的边缘关联概率, 依此对其滤波结果进行加权处理, 最终实现多目标的状态估计与跟踪。

## 2 联合概率数据关联

针对多目标跟踪问题, Bar-shalom等人<sup>[5-6]</sup>提出了基于联合概率数据关联(JPDA)的多目标跟踪方法。假定已知在观测时刻 $t$ 的目标数为 $K$ , 量测数为 $M$ , 建立关联矩阵 $\Omega$ 表征目标与量测的关联关系

$$\Omega = \begin{bmatrix} \Omega_{11} & \dots & \Omega_{1M} \\ \vdots & \ddots & \vdots \\ \Omega_{K1} & \dots & \Omega_{KM} \end{bmatrix}.$$

$\Omega_{kj} = 1$ 表示目标 $k$ 可与量测 $j$ 进行关联, 否则不可关联。目标与量测的关联存在两种策略: 1) 目标分配到量测( $T \rightarrow M$ ), 分配集 $\tilde{\lambda} = \{\tilde{r}, m_C, m_T\}$ ,  $\tilde{r} = \{\tilde{r}_1, \dots, \tilde{r}_k\}$ ,  $\tilde{r}_k = j$ 表示观测 $j$ 来源于目标 $k$ ,  $\tilde{r}_k = 0$ 表示目标 $k$ 未被检测到, 对应的可能关联事件有 $M^K$ 次。2) 量测分配到目标( $M \rightarrow T$ ), 分配集 $\lambda = \{r, m_C, m_T\}$ ,  $r = \{r_1, \dots, r_j, r_j = k\}$ 表示观测 $j$ 来源于目标 $k$ ,  $r_j = 0$ 表示观测 $j$ 来源于杂波, 对应的可能关联事件有 $K^M$ 次。为减少计算量, 根据目标和量测数分别计算两种策略下的关联次数, 选择相应次数最少的关联策略。关联假设集的产生相比于文献[5]确认矩阵的建立和拆分更为容易。考虑到实际场景中通常杂波数大于目标数, 因此适宜采用目标到量测分配的关联策略, 下面简要分析其数据关联和滤波过程。在关联事件 $\tilde{\lambda}$ 下量测的似然为

$$p(y_t|x_t, \tilde{\lambda}) = V^{-M_c} \prod_{k=1}^K p(y_{\tilde{r}_k}|x_k); \quad (1)$$

$$p(y_{\tilde{r}_k}|x_k) = \begin{cases} 1, & \tilde{r}_k = 0; \\ p_T(y_{\tilde{r}_k}|x_k), & \text{else.} \end{cases} \quad (2)$$

当目标数为 $K$ , 可关联的量测数 $m_T$ , 杂波数 $m_C$ 确定时, 如下关联先验成立:

$$p(\tilde{\lambda}_t) = p(m_C) \prod_1^K p(\tilde{r}_k|\tilde{r}_{1:k-1}); \quad (3)$$

$$p(\tilde{r}_k = j|\tilde{r}_{1:k-1}) \propto \begin{cases} 1 - P_D, & j = 0; \\ 0, & j \in \{\tilde{r}_1, \dots, \tilde{r}_{k-1}\}; \\ P_D/M_k, & \text{else.} \end{cases}$$

$M_k = M - \{l: \tilde{r}_l \neq 0, l = 1, 2, \dots, k-1\}$ , 表示能够关联到目标 $k$ 的关联量测数。对每个目标根据联合关联事件建立的量测似然为

$$p_k(y_t|x_t) = \beta_{0k} + \sum_{j=1}^M \beta_{jk} p_T(y_{jt}|x_{k,t}). \quad (4)$$

其中

$$\beta_{jk} = p(\tilde{r}_k = j|y_{1:t}, j = 1, 2, \dots, M) \quad (5)$$

表示目标 $k$ 关联到量测 $j$ 的概率,  $\beta_{0k}$ 表示目标 $k$ 未有有效量测的概率。根据全概率公式

$$\beta_{jk} = p(\tilde{r}_k = j|y_{1:t}) = \sum_{\tilde{\lambda}_t \in \tilde{\Lambda}_t, \tilde{r}_k = j} p(\tilde{\lambda}|y_{1:t}), \quad (6)$$

$\tilde{\Lambda}_t$ 表示所有的联合关联假设集, 有

$$p(\tilde{\lambda}|y_{1:t}) \propto p(\tilde{\lambda}) V^{-M_c} \prod_{j \in \Upsilon} p_{r_j}(y_j|j_{1:t-1}). \quad (7)$$

当目标与量测的关联事件相互独立时, 有

$$\beta_{jk} = p(\tilde{r}_k|y_{1:t}) \sum_{\{\tilde{\lambda}_t \in \tilde{\Lambda}_t, \tilde{r}_k = j\}} \prod_{l=1:k, \tilde{r}_l = j} p(\tilde{r}_l|y_{1:t}) = p(\tilde{r}_k = j|y_{1:t}), \quad (8)$$

此时表征的即是单目标情况下目标与量测的关联概率。目标 $k$ 的状态估计与协方差计算如下<sup>[1]</sup>:

$$\hat{x}^k(t|t) = \sum_{j=0}^M \beta_{jk} \hat{x}_j^k(t|t), \quad (9)$$

$$P^k(t|t) = P^k(t|t-1) - (1 - \beta_{0k}) K^k(t|t) S^{-k}(t|t) K'^k(t|t) + \sum_{j=0}^M \beta_{jk} [\hat{x}_j^k(t|t) \hat{x}_j'^k(t|t) - \hat{x}^k(t|t) \hat{x}'^k(t|t)]. \quad (10)$$

$\hat{x}_j^k(t|t)$ 表示在量测 $j$ 下, 目标 $k$ 轨迹状态的卡尔曼滤波结果,  $K^k(t|t)$ 和 $S^k(t|t)$ 分别表示其增益矩阵和新息协方差矩阵。由式(6)和(8)可知, 为计算目标量测的边缘关联概率, 需要考虑所有可能的联合关联事件。而随着目标数和杂波数的增加, 无论采用何种分配策略, 联合关联事件都会呈现指数增长, 不可避免地引起组合爆炸, 急剧增加计算负荷, 严重降低了系统的实时性能。为此利用马尔可夫链蒙特卡洛方法对其进行改进。

## 3 马尔可夫链蒙特卡洛重要度采样

MCMC作为20世纪最有影响的科学工程成就之一<sup>[7]</sup>, 在高维复杂问题求解、数值计算、概率分布积分以及联合最优化等方面具有独特的优势。MCMC方法建立一个平稳分布的马尔可夫链 $\Psi$ , 通过对其状态空间 $\Omega$ 的采样来实现对其样本概率分布 $\pi(\omega)$ 的逼近。MH<sup>[8]</sup>算法作为MCMC的具体实现形式, 在工程实践中得到了众多应用。MH根据历史样本 $\omega \in \Omega$ 和建议分布函数 $q(\omega, \omega')$ 对下一时刻的样

本  $\omega'$  进行采样, 根据概率  $A(\omega, \omega')$  对样本  $\omega'$  进行接受移动操作, 即

$$A(\omega, \omega') = \min \left\{ 1, \frac{\pi(\omega')q(\omega', \omega)}{\pi(\omega)q(\omega, \omega')} \right\}. \quad (11)$$

当马尔可夫链  $\Psi$  满足不可约分和非周期条件时, 根据遍历性定理  $\Psi$  将最终收敛于  $\pi(\omega)$ . 强杂波下的多目标联合关联概率的计算是一个高维求解的问题, 其状态空间为多目标的联合关联事件集合, 因此可利用 MCMC 方法实现对联合关联概率的求解<sup>[2,9]</sup>. 为提高采样的效率和收敛速度, 利用目标量测关联的概率密度函数进行重要度采样. 基于马尔可夫链蒙特卡洛重要度采样 (MCMCIS) 的多目标跟踪实现步骤如下:

1) 对每一个目标利用波门技术进行可关联判断, 建立全局可关联事件子集

$$E = \{(j, k) | \hat{P}(y(j, t) | y_{1:t-1}^k) \geq \sigma^k, j \leq M, k \leq K\}. \quad (12)$$

其中:  $\hat{P}(y(j, t) | y_{1:t-1}^k)$  表示观测  $j$  落入目标  $k$  预测波门的概率,  $\sigma^k$  为概率门限. 初始化联合关联事件状态采样  $\tilde{\lambda}_0$  和采样次数  $N_s$ .

2) 均匀采样  $u \sim U[0, 1]$ , 如果  $u < 0.05$ , 则  $\tilde{\lambda}' = \tilde{\lambda}$ ; 否则执行 3).

3) 随机选择目标  $k$ , 并对其进行可关联量测重要度采样  $IS\{e = (j, k) \in E\}$ , 计算目标  $k$  波门内量测的关联概率

$$\begin{cases} \eta_0^k \propto \lambda_f(1 - p_d P_G)/p_d, \\ \eta_j^k \propto (2\pi\|S\|^{1/2}) \exp\left\{-\frac{1}{2}\nu_j' S^{-1} \nu_j\right\}. \end{cases} \quad (13)$$

其中:  $\lambda_f$  为杂波密度 (服从均匀泊松分布),  $p_d$  为目标检测概率,  $P_G$  为波门概率,  $\eta_0^k$  为目标漏检 (关联到“0”) 的概率,  $\eta_j^k$  为量测  $j$  来源于当前轨迹的概率,  $\nu_j$  为新息. 利用关联概率  $\eta_j^k$  进行重要度采样, 相应的采样重要度分布为

$$\pi_j^k = \eta_j^k / \sum_j \eta_j^k, j = 0, 1, \dots, m_t^k. \quad (14)$$

$\pi_0^k$  为目标漏检的采样概率,  $\pi_j^k$  为第  $j$  个量测数据被采样的概率,  $m_t^k$  为波门内的备选量测数. 在计算出采样概率后, 以概率  $\pi_j^k$  对量测  $j$  进行采样.

4) 如果  $e \in \tilde{\lambda}$ , 则进行删除操作  $\tilde{\lambda}' = \tilde{\lambda} - e$ , 对目标  $k$  进行重要度采样; 否则执行 5).

5) 如果  $\tilde{\lambda}$  中目标  $k$  没有有效量测且量测  $j$  来源于杂波, 则进行增加操作: 将量测  $j$  关联到目标  $k$ , 即  $\tilde{\lambda}' = \tilde{\lambda} + e$ ; 否则执行 6).

6) 如果量测  $j$  已被关联且目标  $k$  没有有效量测, 则进行交互操作: 将量测  $j$  关联到目标  $k$ , 并将量测  $j$  的原有目标关联目标  $k$  的原有量测; 否则  $\tilde{\lambda}' = \tilde{\lambda}$ .

7) 根据 MH 算法对关联事件  $\tilde{\lambda}'$  按照概率  $\alpha$  进行

接受, 即

$$\alpha = \min\{p(\tilde{\lambda}' | y_{1:t})/p(\tilde{\lambda} | y_{1:t}), 1\}, \quad (15)$$

$p(\tilde{\lambda} | y_{1:t})$  的计算如式 (7). 重复 2)~6), 直至采样结束.

8) 根据联合状态事件采样的结果, 按照下式统计目标量测的边缘关联概率:

$$\hat{\beta}_j^k = \frac{1}{N_s} \sum_{i=1}^{N_s} I((j, k) \in \lambda_t), \quad (16)$$

其中  $I$  为指示函数.

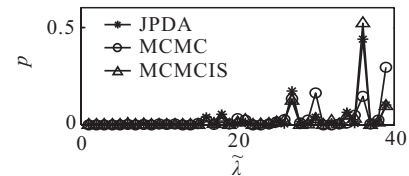
9) 根据式 (9) 和 (10) 进行滤波与目标状态估计.

上述构建的马尔可夫链满足不可约分和非周期性条件, 根据遍历性原理最终将收敛于联合关联概率. 因无须对联合关联事件进行枚举, 故避免了组合爆炸问题.

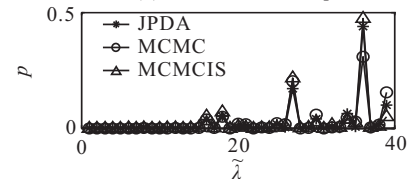
## 4 仿真实验与分析

### 4.1 联合关联概率计算

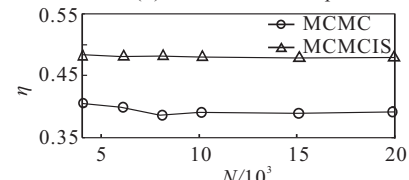
首先采用 JPDA, MCMC 和 MCMCIS 三种方法进行联合关联概率计算仿真对比实验. 定义接受效率  $\eta$  为马尔可夫链蒙特卡洛过程中接受次数与采样总数的比值,  $\sigma_{\Delta p}$  为求解的联合关联事件  $\tilde{\lambda}$  的概率误差



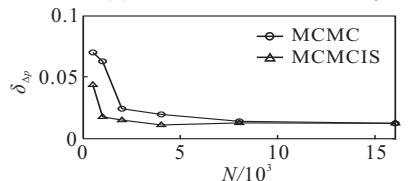
(a) 1000次采样  $\tilde{\lambda}$  及  $p$



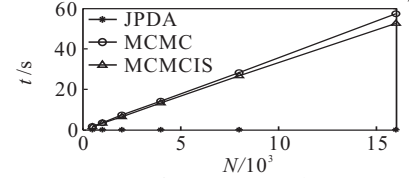
(b) 2000次采样  $\tilde{\lambda}$  及  $p$



(c) 采样次数  $N$  与接收效率  $\eta$



(d) 采样次数  $N$  与关联概率误差标准差  $\delta_p$



(e) 采样次数  $N$  与计算时间  $t$

图 1 低密度杂波多目标联合关联概率计算

$\Delta p$  的方差.  $\eta$  越大, 表明采样过程越能更快地收敛于  $\pi(\omega)$ ;  $\sigma_{\Delta p}$  越小, 表明求解的联合关联事件概率越逼近真实值. 建立低密度和高密度杂波两个仿真场景, 杂波密度参数分别为 0.06 和 0.5, 预测的目标位置为  $[0, 1]^T$ ,  $[0, -1]^T$ ,  $[1, 1]^T$ ,  $[1, -1]^T$ , 测量噪声协方差为 1, 目标检测概率为 0.98, 波门关联概率为 0.95, 杂波在观测区域内均匀分布, 杂波数目服从泊松分布. 仿真运行环境: Matlab, 9, Core2 3.0G. 低密度杂波场景的可行关联事件数为 39, 在不同采样次数下的仿真结果如图 1 所示.

图 1 的仿真结果表明, 在低密度杂波环境下 MCMC 和 MCMCIS 方法能够在较少采样次数下实现对联合关联概率的求解, 其精度随着采样次数的增加而提高, 最终收敛到 JPDA 的求解结果. MCMCIS 方法相比于 MCMC 方法具有更高的接受效率, 表明能更快地逼近真实联合关联概率. 同时, 对于具有较大概率的真实联合关联事件, MCMCIS 方法能够获得更大的概率值. 其原因在于重要度采样技术使得马尔

可夫链向着高概率密度联合事件状态转移, 从而对其进行更多采样, 提高了接受效率. 但是由于关联事件非常少, 相比于 JPDA 方法反而需要花费更多的计算资源.

高密度杂波场景的可行关联事件数为 82 012, 考察目标量测的边缘关联概率, 不同采样次数下的仿真结果如图 2 所示. 图 2 的仿真结果表明, 在高密度杂波环境下由于联合关联事件巨大导致 JPDA 出现组合爆炸, 边缘关联概率计算时间达到 160 s, 严重影响了系统的实时性能. MCMC 与 MCMCIS 方法只需要 4 000 次采样即可获得比较接近真实值的边缘关联概率, 其计算时间 ( $\approx 16$  s) 远远小于 JPDA 方法, 极大地改善了系统的实时性能. 随着采样次数的增加, 计算时间呈线性增长. MCMCIS 方法相比于 MCMC 方法概率权值更为集中, 有利于降低边缘杂波的干扰, 提高真实目标量测的联合关联概率, 改善了跟踪精度.

#### 4.2 多目标轨迹跟踪

进一步应用 JPDA, MCMC 及 MCMCIS 方法进行多目标轨迹跟踪与状态估计仿真对比实验. 场景参数: 杂波密度为 0.01, 检测概率为 0.999, 波门关联概率为 0.9999, 测量噪声协方差为 25, 跟踪波门内平均杂波数为 15, 观测周期为 5 s, 累积观测 50 次. 目标进行匀速直线运动, 初始状态 (位置及速度) 为:  $[-40, 350, 1, -1.6; -50, 0, 1, 0; -40, -400, 1, 3.1]^T$ , 50 次蒙特卡洛仿真结果如表 1 和图 3 所示.

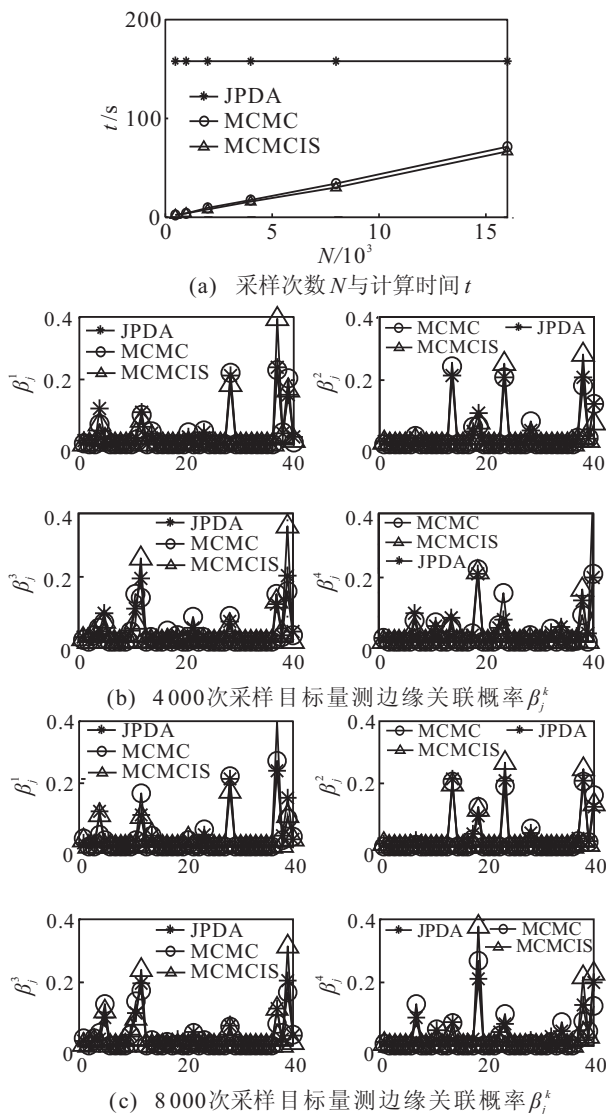


图 2 高密度杂波多目标联合关联概率计算

表 1 50 次蒙特卡洛仿真 RMSE 误差

	Target#1	Target#2	Target#3
JPDA	4.222 7	4.538 7	4.157 7
MCMC	4.578 6	4.708 8	4.498 0
MCMCIS	4.255 1	4.648 3	4.283 1

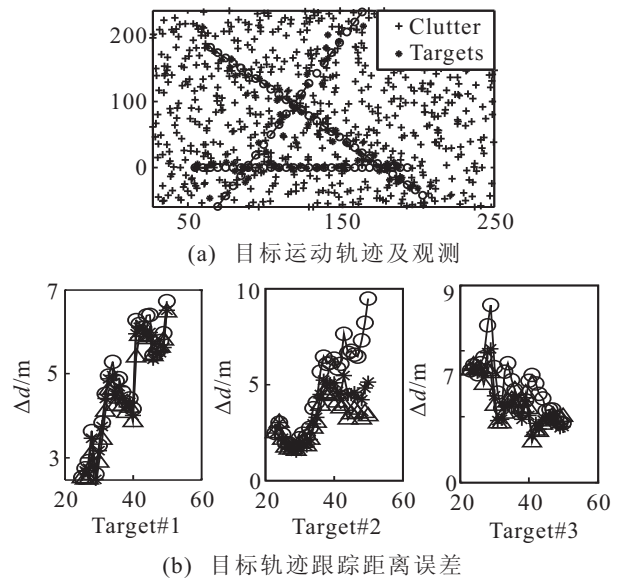


图 3 高密度杂波下多目标轨迹跟踪

上述仿真结果表明,算法 MCMC 和 CMCIS 在高密度杂波下能较好地跟踪目标轨迹. 由于引入了重要度采样, MCMCIS 相对 MCMC 具有更好的跟踪精度. 在目标轨迹生命周期中, 邻近目标的干扰非持续存在, 在跟踪过程中对于不存在多目标与量测相互干扰的时刻将直接采用单目标概率数据关联 (PDA) 进行处理, 在目标发生交叉区域利用 JPDA 和 MCMC 及 MCMCIS 进行处理.

## 5 结 论

针对高密度杂波下的多目标跟踪问题, 本文提出了一种基于马尔可夫链蒙特卡罗重要度采样的多目标跟踪方法, 有效地解决了联合概率数据关联中的“组合爆炸”问题, 并且改善了采样效率. 仿真实验表明, 该方法在保持目标轨迹跟踪精度的同时有效地降低了计算负荷, 显著提高了系统的实时性能, 具有较好的工程应用价值.

## 参考文献(References)

- [1] 何友, 王国宏, 彭应宁, 等. 多传感器信息融合及应用[M]. 北京: 电子工业出版社, 2001.  
(He Y, Wang G H, Peng Y N, et al. Multisensor information fusion with applications[M]. Beijing: Publishing House of Electronics Industry, 2001.)

- [2] Oh S, Russell S, Sastry S. Markov chain Monte Carlo data association for multi-target tracking[J]. IEEE Trans on Decision and Control, 2009, 54(3): 481-497.
- [3] Reid D. An algorithm for tracking multiple targets[J]. IEEE Trans on Automatic and Control, 1979, 24(6): 843-854.
- [4] Cox S Hingorani. An efficient implementation of Reids multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking[J]. IEEE Trans on Pattern and Machine Intelligence, 1996, 18(2): 130-150.
- [5] Bar-Shalom Y, Fortman T. Tracking and data association[M]. San Diego: Academic press, 1988.
- [6] Vermaak J, Godsill S J, Perez P. Monte Carlo filtering for multi-target tracking and data association[J]. IEEE Trans on Aerospace and Electronic Systems, 2005, 41(1): 309-332.
- [7] Beichl I, Sullivan F. The metropolis algorithm[J]. Computing in Science and Engineering, 2000, 2(1): 65-69.
- [8] Hastings W K. Monte Carlo sampling methods using Markov chains and their applications[J]. Biometrika, 1970, 57(1): 97-109.
- [9] Kan Z, Balch T, Dellaert F. MCMC-based particle filtering for tracking a variable number of interacting targets[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(11): 1805-1918.

(上接第1401页)

- [3] Pavlov I P. Conditioned reflexes[M]. Oxford: Oxford University Press, 1927.
- [4] Brembs B. Research: Neurobiology of behavior[EB/OL]. (2010-06-01). <http://brembs.net>.
- [5] Brembs B, Plendl W. Double dissociation of PKC and AC manipulations on operant and classical learning in drosophila[J]. Current Biology, 2008, 18(15): 1168-1117.
- [6] Brembs B. The importance of being active[J]. J of Neurogenetics, 2009: 23(1/2): 120-126.
- [7] Zalama E, Gomez J, Paul M, et al. Adaptive behavior navigation of a mobile robot[J]. IEEE Trans on Systems, Man, and Cybernetics, Part A: Systems and Humans, 2002, 32(1): 160-169.
- [8] Itoh K, Miwa H, Matsumoto M, et al. Behavior model of humanoid robots based on operant conditioning[C].

IEEE/RAS Int Conf on Humanoid Robots. Piscataway: IEEE, 2005: 220-225.

- [9] 王雪松, 程玉虎, 易建强, 等. 基于 Elman 网络的非线性系统增强式学习控制[J]. 中国矿业大学学报, 2006, 35(5): 653-657.  
(Wang X S, Cheng Y H, Yi J Q, et al. Reinforcement learning control for nonlinear systems based on Elman neural network[J]. J of China University of Mining and Technology, 2006, 35(5): 653-657.)
- [10] Guo M, Liu Y, Malec J. A new Q-learning algorithm based on the metropolis criterion[J]. IEEE Trans on System, Man, and Cybernetics, Part B: Cybernetics, 2004, 34(5): 2140-2143.
- [11] Obayashi M, Nakahara N, Kuremoto T, et al. A robust reinforcement learning using the concept of sliding mode control[J]. Artif Life Robotics, 2009, 13(2): 526-530.