

基于 K -Medoids 聚类的多传感器航迹关联算法

徐 丽^{1,2}, 马培军¹, 苏小红¹

(1. 哈尔滨工业大学 计算机科学与技术学院, 150001 哈尔滨, xuli_hit@126.com; 2. 哈尔滨工程大学
计算机科学与技术学院, 150001 哈尔滨)

摘 要: 为有效解决目标密集环境下的航迹关联问题, 提出了一种基于 K -Medoids 聚类的航迹关联算法. 该算法采用局部航迹与系统航迹进行关联的策略, 将系统航迹作为 Medoids, 降低了需要关联的航迹对数量, 避免了 K -Medoids 的固有缺陷, 很大程度上提高了关联算法的效率. 通过采用无穷范数计算采样点点迹距离求出了两条航迹的近似距离, 这使得关联判决能考虑历史和当前航迹, 提高了正确关联率. 在多传感器多目标环境下讨论了其具体实现过程, 仿真实验结果验证了该算法的有效性和优越性. 该算法在存在噪音和离群点时, 具有很强的健壮性, 适合目标密集环境.

关键词: 航迹关联; K -Medoids 聚类; 无穷范数; 多传感器; 多目标

中图分类号: TP391 **文献标志码:** A **文章编号:** 0367-6234(2012)01-0107-04

A multi-sensor track association algorithm based on K -Medoids clustering

XU Li^{1,2}, MA Pei-jun¹, SU Xiao-hong¹

(1. School of Computer Science and Technology, Harbin Institute of Technology, 150001 Harbin, China, xuli_hit@126.com;
2. College of Computer Science and Technology, Harbin Engineering University, 150001 Harbin, China)

Abstract: To solve the problem of track association in dense target environment, a new track association algorithm based on K -Medoids clustering is presented. The algorithm makes the sensor tracks associate with the system tracks and selects the system tracks as Medoids, which greatly reduce the number of track pairs needed to associate, avoid the inherent defects of K -Medoids clustering, and improve the efficiency of the algorithm. In addition, the approximate distance between the two tracks is acquired by calculating the distance at a sampling point with the method of infinite norm, and then the track pair with minimum distance is selected as the associated tracks. The consideration of the historical and current track improves the probability of correct association. Finally, the algorithm is implemented in the multi-sensor and multi-target environment and the simulation results prove the effectiveness and superiority of the algorithm. The algorithm has a strong robustness in the presence of noise and outliers, and is suitable for the dense target environments.

Key words: track association; K -Medoids clustering; infinite norm; multi-sensor; multi-target

航迹融合必须以航迹关联为前提, 因此关联判定的准确性将直接影响到整个航迹融合系统的性能. 航迹关联备受国内外学者关注, 一直是研究的热点问题之一. Ashraf M. Aziz^[1] 提出了一种基

于模糊均值聚类的航迹融合方法来解决分布式多传感器多目标多属性重叠覆盖场景中的航迹关联和融合问题. Baoguo Tian 等^[2] 将航迹关联问题转化成多维分配问题进行求解. Songhwei Oh 等^[3] 用马尔科夫链蒙特卡洛数据关联算法较好地解决了目标密集环境下的航迹关联问题. Huang Youpeng 等^[4] 提出了一种基于灰色关联分析的航迹关联算法, 有效地实现了异构传感器的航迹关联. Kusha Panta 等^[5] 提出了混合高斯概率假设密度

收稿日期: 2011-01-20.

基金项目: 国家自然科学基金资助项目(60773067); 中央高校基本科研业务费专项资金资助项目(HEUCF100604).

作者简介: 徐 丽(1980—), 女, 博士研究生, 讲师;
马培军(1963—), 男, 教授, 博士生导师;
苏小红(1966—), 女, 教授, 博士生导师.

滤波的航迹关联. Mei Duan 等^[6]提出了基于神经网络的航迹关联算法. 这些算法的提出提高了多传感器航迹融合系统的效率. 近年来, 还有很多学者从其他角度研究了航迹关联问题. Donald E. Maurer^[7]对关联的信息进行处理, 提高了整个系统的效率. Lance M. Kaplan 等^[8]提出了关联多条航迹的代价函数. Dimitri J. Papageorgiou 等^[9]扩展了最佳偏差关联假设和相应的偏差关联似然函数, 使其更适合系统级航迹歧义管理. 这些成果很大程度上促进了航迹关联问题的研究. 特别是, 文献^[10]根据数据列因素之间发展态势的相似或相异程度来衡量航迹间接接近的程度, 首次用模式识别的方法解决了航迹关联问题, 为求解航迹关联问题探索了一条新的途径.

本文提出了一种新的基于 *K*-Medoids 聚类的航迹关联算法. 该算法采用局部航迹与系统航迹关联的策略, 大大降低了需要关联的航迹对数量, 从而提高了关联算法的效率.

1 系统描述

假设 *M* 个传感器观测杂波中的 *T* 个目标, 在固定时间间隔内获取观测, 每个观测都由几个量测组成, 共观测 *n* 步.

设 $X^i(k)$ ($1 \leq i \leq T$) 为第 *k* 个测量时刻目标 *i* 的状态向量, 假设输入项为零, 则目标运动模型为

$$X^i(k+1) = F(k)X^i(k) + G(k)u^i(k) + v^i(k), \quad k = 1, 2, \dots, n; i = 1, 2, \dots, T. \quad (1)$$

式中: $X^i(k+1)$ 为 *k+1* 时刻目标 *i* 的状态向量; $F(k)$ 为状态转移矩阵; $G(k)$ 为输入控制矩阵; $u^i(k)$ 为加速度输入矩阵; $v^i(k)$ 为离散时间白噪声序列, 满足

$$E[v^i(k)] = 0, \quad (2)$$

$$E[v^i(k)v^i(k)^T] = Q^i(k). \quad (3)$$

测量方程可表示为

$$Z_l(k) = H(k)X^i(k) + w_l(k). \quad (4)$$

式中: $w_l(k)$ 为零均值, 方差为 $R_l(k)$ 的 Gauss 观测噪声; $H(k)$ 是观测矩阵.

2 基于 *K*-Medoids 聚类的航迹关联算法

在分布式结构的航迹融合系统中, 每个传感器都独立地处理它的局部量测, 产生局部航迹并送至融合中心, 融合中心根据各传感器的航迹数据完成航迹关联和航迹状态估计融合, 形成全局估计. 航迹关联的工作就是将来自不同传感器的航迹进行分组, 使得同一组的航迹代表同一个目标. 而聚类是将要处理的对象分成若干个类, 使得

同一类的对象尽量相似, 不同类的对象尽量相异. 二者都可用于发现隐藏在数据背后的分组和数据分布信息. 实际上, 二者的目标和作用是一致的. 通过上述分析, 可以从一个新的角度重新理解航迹关联问题.

2.1 关联策略

K-Medoids 聚类算法属于划分方法中的一种常用的聚类算法, 具有较高的准确性. 因为 Medoids 不容易被极端数据影响, 当存在噪声和孤立点数据时, *K*-Medoids 算法依然很健壮, 适合数据比较密集的数据集. 然而, *K*-Medoids 算法也存在一些缺陷: 1) 初始化敏感; 2) 聚类结果多样化; 3) 在进行 Medoids 轮换时需遍历所有非 Medoids, 执行代价较高.

由于在分布式多传感器航迹融合系统的实际应用中, 所使用的传感器和每个传感器所扫描的目标数量都很多, 如果对所有来自不同传感器的航迹都进行两两关联及融合处理, 那将会给系统带来沉重的负担. 因此, 本文采用将局部航迹与系统航迹进行关联的策略, 指定每条系统航迹作为一个固定的 Mediod, 这样就避免了 *K*-Medoids 算法初始化的随机性, 也避免了 Medoids 轮换的沉重代价. 而这种策略, 本身就大大降低了需要关联的航迹对数量, 使系统的效率得到了很大程度的提高. 另外, 由于来自同一传感器的航迹都是由不同目标形成的, 所以一条系统航迹最多只能与来自某个传感器的一条局部航迹关联成功; 又由于每条系统航迹都源于不同的目标, 所以来自一个传感器的某条局部航迹最多只能与一条系统航迹关联成功, 因此算法将根据规则最多为一条局部航迹找到一条系统航迹与其关联, 这样避免了聚类结果的多样性.

2.2 算法描述

Step1 确定 Mediods.

设系统航迹和来自传感器 *l* 的局部航迹的航迹号集合分别为

$$U_1 = \{1, 2, \dots, n_1\}, \quad (5)$$

$$U_2 = \{1, 2, \dots, n_2\}. \quad (6)$$

即系统航迹集合中有 n_1 条系统航迹, 来自传感器 *l* 的局部航迹集合中有 n_2 条局部航迹.

指定每条系统航迹作为一个 Mediod, 每个 Mediod 标识一个类, 共 n_1 个类. 记为

$$x_s = \{x_{s1}, x_{s2}, \dots, x_{sn_1}\}. \quad (7)$$

传感器 *l* 的 n_2 条局部航迹记为

$$x_l = \{x_{l1}, x_{l2}, \dots, x_{ln_2}\}. \quad (8)$$

Step2 计算两条航迹 *k* 时刻点迹的距离.

假设航迹 i 在 k 时刻经过标准化处理后的状态向量为

$$\mathbf{x}_i(k) = [r_{i1}(k), r_{i2}(k), \dots, r_{in}(k)]', \quad (9)$$

相应的分辨率为

$$\mathbf{\Delta}_i = [\delta_{i1}, \delta_{i2}, \dots, \delta_{in}]'. \quad (10)$$

式中: r_k 为航迹的特征; n 为特征的个数; δ_k 为每个特征相对应的分辨率。

利用无穷范数定义航迹 i 和航迹 j 在 k 时刻的点迹距离为

$$d_{ij} = \begin{cases} \|\mathbf{x}_j(k) - \mathbf{x}_i(k)\|_{\infty}, & i \neq j; \\ \|\mathbf{\Delta}_i\|, & i = j. \end{cases} \quad (11)$$

即

$$\begin{cases} d_{ij} = \max_{1 \leq l \leq n} |r_{il}(k) - r_{jl}(k)|, \\ d_{ii} = \max_{1 \leq l \leq n} |\delta_{il}|. \end{cases} \quad (12)$$

进而, 利用平均值法计算出描述两条航迹接近程度的近似距离, 即

$$d_{ij} = \frac{1}{n} \sum_{k=1}^n d_{ij}(k). \quad (13)$$

Step 3 构造来自传感器 l 的局部航迹与 Medoids 的距离矩阵。

当计算出两条航迹的近似距离之后, 可以构造出来自传感器 l 的局部航迹与 Medoids 的距离矩阵为

$$\mathbf{d} = \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n_2} \\ d_{21} & d_{22} & \dots & d_{2n_2} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n_1 1} & d_{n_1 2} & \dots & d_{n_1 n_2} \end{bmatrix}.$$

Step 4 关联准则。

根据与 Medoids 距离最近的原则, 即 $d_i = \min_j d_{ij}$ 且 $d_i < d_{ii}$. 将待关联的局部航迹分到各个类中去, 即指派每个局部航迹给离它最近的 Medoid 所代表的类. 当一个 Medoid 与多条局部航迹距离相等且都为最小时, 由于一条系统航迹最多只能与来自某个传感器的一条局部航迹来自同一个目标, 且目标位置是关联判决中一个较为关键的属性, 算法将选择与该 Medoid 平均位置无穷范数最小的局部航迹作为最终的关联航迹. 当 $d_i > d_{ii}$ 时, 说明该局部航迹不与任何 Medoids 关联, 这时指定此局部航迹作为一个新的 Medoid. 这样就给出了系统航迹与来自传感器 l 的局部航迹的关联判决。

Step 5 多义性处理。

一条来自某个传感器的局部航迹最多只能与一条系统航迹关联. 因此当一条局部航迹与多条系统航迹关联成功时, 需要进行多义性处理. 在这

种情况下算法选择与该局部航迹距离最小的系统航迹为关联航迹, 如果与该局部航迹距离最小的系统航迹依旧有多条, 则算法选取其中平均位置无穷范数最小的那条系统航迹作为最终与该局部航迹关联的航迹. 经过多义性处理后, 每条局部航迹就最多只能与一条系统航迹关联成功。

经航迹关联后, 将关联成功的航迹对进行航迹状态估计融合。

3 实验分析

3.1 初始设置

为了讨论问题方便, 假设送至融合中心的所有状态估计都在相同的坐标系里, 并且各传感器同步采样, 数据的传输延迟时间为 0. 仿真实验采用两个传感器同时观测目标, 模拟目标在三维空间中做变速机动运动. 为了验证算法的性能, 采用 Monte Carlo 方法对本文算法进行 50 次仿真. 仿真考虑两种情况: 1) 模拟中等密度目标环境, 开始进入公共区的目标为 60 批; 2) 模拟密集目标环境, 开始进入公共区的目标为 120 批. 图 1 和图 2 分别给出了在公共观测区域 60 批目标和 120 批目标的运动轨迹。

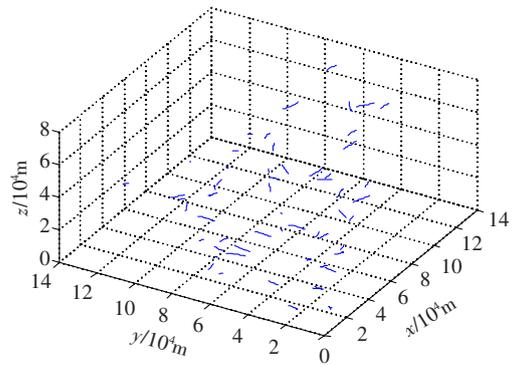


图 1 60 批目标航迹图

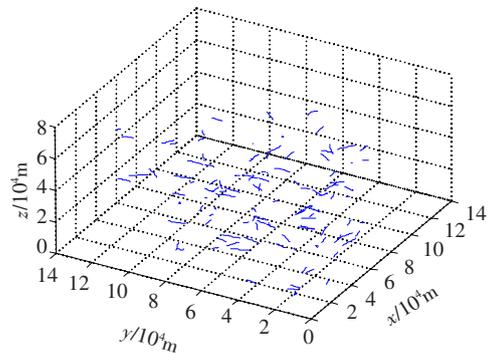


图 2 120 批目标航迹图

3.2 实验结果与分析

图 3 与图 4 分别给出了在 60 批目标下分别采用最近邻域法 (NN)、模糊 C 均值法 (FCM) 及本文算法 (*K*-Mediods), 仿真 50 次后的平均正确关联率 (*P_c*) 曲线和平均错误关联率 (*P_e*) 曲线。

图 5 与图 6 分别给出了在 120 批目标下分别采用最近邻域法 (NN)、模糊 C 均值法 (FCM) 及本文算法 (*K*-Mediods), 仿真 50 次后的平均正确关联率 (*P_c*) 曲线和平均错误关联率 (*P_e*) 曲线。

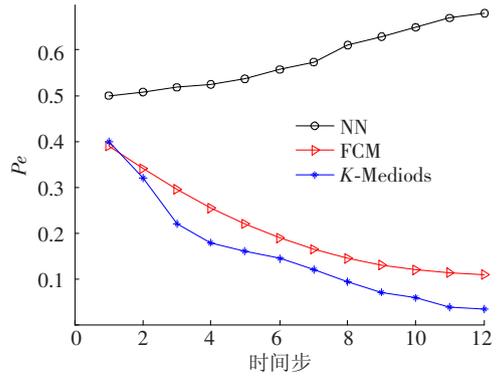


图 6 120 批目标下错误关联率对比

表 1 统计了采用最近邻域法 (NN)、模糊 C 均值 (FCM) 及本文算法 (*K*-Mediods) 对各目标进行航迹关联的平均计算时间。

关联算法	60 批目标下平均 关联时间	120 批目标下平均 关联时间
NN	14	53
FCM	75	382
<i>K</i> -Mediods	61	272

从实验结果可以看出最近邻域法速度很快, 但是正确关联率最低. 该算法不适合中等密度和密集目标环境, 特别在密集目标环境下, 随着关联时间步的增加, 其正确关联率逐步下降. 模糊 C 均值法的正确关联率较高, 在中等密度目标环境下与 *K*-Mediods 方法相当, 但在密集目标环境下, 其正确关联率与 *K*-Mediods 方法的差距变大, 且效率较低. 本文算法 *K*-Mediods 方法的正确关联率在中等密度和密集目标环境下都为最高, 在目标密集环境下, 仍然可以达到较高的正确关联率.

4 结 论

1) 基于 *K*-Medoids 聚类的航迹关联算法采用局部航迹与系统航迹关联的策略, 减少了需要关联的航迹对数量, 缩短了关联时间, 提高了系统的效率; 系统航迹被指定为 Mediods, 使得用 *K*-Mediods 聚类方法求解航迹关联问题克服了该方法本身的缺陷.

2) 通过用无穷范数量化两条航迹在采样点的点迹距离得到了两条航迹的近似距离, 从而确定了来自某个传感器的局部航迹与系统航迹的关联矩阵, 使得关联判决能考虑当前和历史航迹, 提高了正确关联率.

(下转第 148 页)

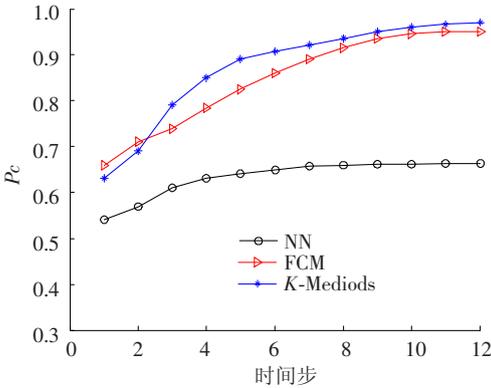


图 3 60 批目标下正确关联率对比

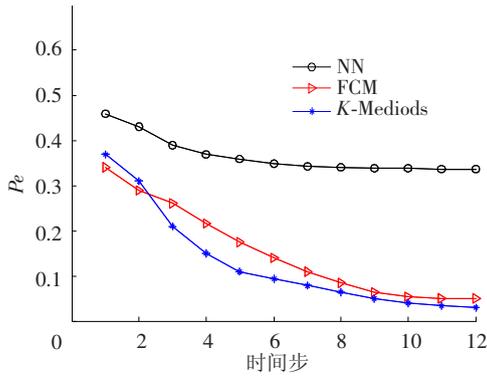


图 4 60 批目标下错误关联率对比

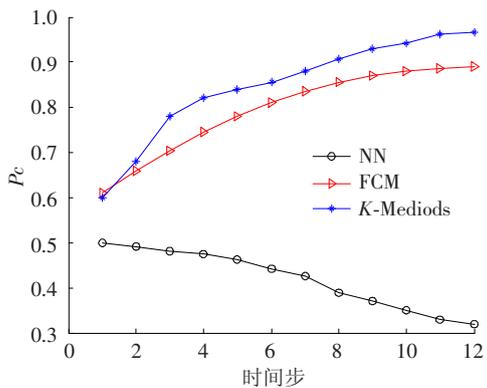


图 5 120 批目标下正确关联率对比