

网格聚类在多雷达数据融合算法中的应用

舒红平^{1,2}, 徐振明², 邹书蓉², 何嘉²

(1. 四川大学计算机学院 成都 610064; 2. 成都信息工程学院计算机系 成都 610041)

【摘要】应用网格聚类的方法区分同一雷达接收的不同目标的观测数据,通过类间数据融合,实现同一目标不同雷达接收数据的融合,以便对多目标进行实时跟踪。研究了观测数据网格聚类的基本思想、形式化描述及算法实现,讨论了对机动目标跟踪的Kalman滤波方程及空管系统中易于计算的各参数矩阵理论依据及相应的初值。仿真结果表明,通过网格聚类能很好地区分不同目标,聚类后再进行跟踪融合更加准确。

关键词 空中交通管理; 数据融合; 网格聚类; 雷达
中图分类号 V21 文献标识码 A

Research on Multiradar Data Fusion Algorithm Based on Grid Clustering

SHU Hong-ping^{1,2}, XU Zheng-ming², ZOU Shu-rong², HE Jia²

(1. College of Computer Science, Sichuan University Chengdu 610064;
2. Department of Computers, Chengdu University of Information Technology Chengdu 610041)

Abstract This paper studies application grid clustering method to distinguish the target observation data which is received by the same radar. The fusion of the same goal's observation data received by different radars is realized through the integration of different observation data, so as to realize real-time tracking of multiple targets. The basic thought of grid observation data clustering, and the formal description algorithm are studied. The Kalman filter equations for maneuvering target tracking are described. Parameter matrix theoretical basis for the simplified calculation and the corresponding initial matrix are given for air traffic control system. The simulation result indicates that the grid clustering is able to discriminate different targets well and to carry on the track fusion more accurate.

Key words air traffic control; data fusion; grid clustering; radar

多雷达数据融合处理是整个雷达信息处理的中心,它把多通道雷达接口处理单元送来的多路雷达信息进行实时融合处理成综合雷达数据,并将处理结果送往冲突、低高、区域预警系统及有关管制员席位作进一步处理。

多部雷达同时监控某一区域的多批目标时,由于通信延迟,雷达天线扫描周期不同及各雷达系统误差等多种原因,在某一时刻情报融合中心接收到的多部雷达测量数据呈现团状,且分布在目标真实值的周围。如何做到既要区分同一传感器观测的不同目标,又要关联不同传感器观测的同一目标,就成为数据融合中目标观测数据关联的难点。当目标数量少且相互位置较远时,关联容易实现;而当目标数量多且相互位置较近时,由于信息处理中心接

收到的同一传感器对不同目标的观测数据可能比不同传感器对同一目标的观测数据相似度更近,关联就容易产生错误。为避免产生错误的关联结果,本文提出通过网格聚类分析方法对目标观测数据进行聚类,通过对每一时刻传感器数据进行聚类,以区分源于不同目标的数据;通过类间数据融合,实现对多目标的实时跟踪。

1 测量数据的网格聚类

聚类需要解决的问题是将已给定的若干无标记的模式聚集起来使之成为有价值的类^[1]。基于划分的聚类分析算法是最常用的聚类分析算法之一,包括K-平均值和K-中心点的聚类算法。K-平均值算法最早是由文献[1]提出来的,在该算法中,每个类用

收稿日期: 2007-09-14

基金项目: 国家自然科学基金(60572175); 四川省青年科技基金(06ZQ026-054)

作者简介: 舒红平(1974-), 男, 博士生, 副教授, 主要从事数据融合、数据挖掘及智能决策等方面的研究。

该类中现有对象的平均值表示。该算法对异常数据非常敏感,文献[2]提出了基于K-中心点的PAM和CLARA算法。这两种算法不是用样本的平均值表示,而是用接近聚类中心的中心点来表示。K-中心点算法对异常数据不敏感,但计算量比K-平均值算法要大,一般情况下只适用于小数据量。

雷达探测数据的特征使选择聚类算法时必须满足以下条件:

(1) 目标的类别数未知。因为即使在初始聚类完成以后,由于误警和漏警的存在,目标可能出现分批和合批的情况。

(2) 较强的鲁棒性。个别误警和漏警的存在不会影响分类的结果。

(3) 算法运行效率高。因为雷达探测的数据是实时的,每个扫描周期都会产生一批数据,观测数据量大。

无论是K-平均值算法还是K-中心点算法都存在如下不足^[4]:

(1) 必须事先假定类别个数,即K值,但进行雷达数据的聚类处理时,K值很难事先确定。

(2) 初始聚类中心的随机选取可能会陷入局部最优解,而难以获得全局最优解。

(3) 难以消除噪声的影响。

基于以上的考虑,可采用网格聚类分析方法来实现目标观测数据的聚类以解决上述问题。网格聚类算法不仅克服了K-平均值算法和K-中心点聚类算法的不足,而且提高了算法的运行效率。

聚类之前,先假定所有雷达测量数据已经被校准过,所有传感器的扫描周期相同,如果不同,需进行时空对准处理,且要统一转换到一个统一的坐标系下,故需进行坐标变换处理。

1.1 算法基本思想

对多维空间的网格,将每一维划分成区间,对网络进行编码并统计每个网格中的记录个数,每个网格用网格中心点作为代表点。判断每个网格中的记录是否小于异常点阈值,若小于,则标记该网格为异常点,否则为非异常点。找出距离最远的两个非异常点网格分别代表两个初始类,判断其他未被分类的非异常点网格与距离最近的现有类的代表网格之间的距离是否小于聚类阈值,若小于,则将该网格分配到对应类中;否则将该网格标记为一新类。一直迭代直至所有网格分类完毕。

雷达观测数据的网格分布示意图如图1所示。

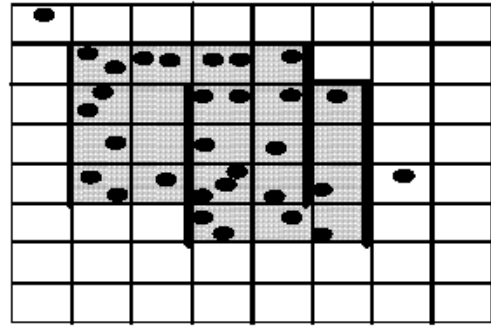


图1 雷达观测数据的网格分布示意图

1.2 算法的形式化描述

定义 1 设 $A = \{D_1, D_2, \dots, D_n\}$ 是 n 个有界定义域, $S = D_1 \times D_2 \times \dots \times D_n$ 就是一个 n 维空间,称 $D_1 \times D_2 \times \dots \times D_n$ 是 S 的维^[3]。

定义 2 (t 区间划分)对 n 维空间 S 的每一维分成相同的 t 个区间,整个空间被分成了有限个不相交的类(矩形单元),各矩形单元的尺度依据空域划分情况来确定,每一个这样的矩形单元可以描述为 $\{u_1, u_2, \dots, u_n\}$ 。

定义 3 (类或矩形单元)设 G 为一个区间的元素的集合,它共有 n 个元素,记为 $g_i, i = 1, 2, \dots, n$,给定一阈值 $H > 0$,对 G 中任意一个元素 g_i ,总存在另一个元素 g_j ,它们的距离(欧几里德距离)不大于阈值 H ,即有 $d_{ij} \leq H$,则称 G 为类或矩形单元。

定义 4 (网格编码)设每一维的空间的元素用 $i \times j$ 行列矩形网格来表示,网格所在的行列位置序号即为该网格的编码。

定义 5 (相邻网格)^[4]两个网格 u_i 和 u_j ,若它们之间有一个公共面,且所在行列序号要么行序号相等,要么列序号相等,则称 u_i 和 u_j 为相邻网格。

1.3 算法实现

基于网格的雷达数据聚类算法分为6步,具体描述如下:

Input: 雷达观测数据(聚类对象)异常点阈值 D_c , 聚类阈值 D_j ;

Output: 区域表、类_记录数表(用于存放分类的个数和每类的记录总和);

Step1: 将定义1多维网格的每一维划分为区间,实现多维空间的网格化,每个网格划分按10 km进行。并按定义4进行网格编码,统计每个网格中的记录个数。

Step2: If 每个网格的记录 $< D_c$ then
将该网格标记为异常点;

else

将该网格标记为非异常点。

Step3 : 对每个非异常点的网格, 计算其与其他网格之间的欧几里德距离 $d(i, j)$, $d(i, j)$ 表示第 i 个非异常点网格与第 j 个非异常点网格之间的距离。

Step4 : If $d(i, j) = \text{MAX}$ then 第 i 个非异常点网格与第 j 个非异常点网格分别代表两个初始类。

Step5 : If 未被分类的非异常点网格与距离最近的现有类的代表网格之间的距离 $< D_j$

then 将该网格分配到对应类中;

Else 将该网格标记为一新类。

递归调用本步, 直至所有网格分类完毕。

2 数据融合和目标跟踪

2.1 雷达观测数据融合

飞机在空中飞行时可以设定它是等速运动的, 由机动和大气湍流引起的加速可以看作对等速运动的一种摄动。机动的快速程度用 $\alpha = \frac{1}{T_c}$ 来反映, T_c 是

相关函数的时间常数。对于快速程度不同的机动, α 取值不同。在一个 T_c 时间内, 可得到 N 个观测值构成的一个类 $C = \{Z^i(t_c) | i = 1, 2, \dots, N\}$ 。为了充分利用各个雷达目标发现的信息, 以便得到更加精确的目标航迹, 应对这 N 个测量进行融合。采用状态方程描述目标的运动特性为:

$$\begin{cases} X^i(t_{k+1}) = \Phi(T, \alpha)X^i(t_k) + U^i(t_k) \\ Z^i(t_k) = HX^i(t_k) + W^i(t_k) \end{cases} \quad (1)$$

式中 $U^i(t_k)$ 为非齐次驱动输入, 是一种离散的白噪声矢量序列; $W^i(t_k)$ 为 t_k 时刻的观测噪声。

$$E[U^i(t_k)] = 0, \quad E[W^i(t_k)] = 0$$

$$E[W^i(t_k)W^T(t_k)] = R_i = \sigma_w^2(i)I, \quad E[U^i(t_k)U^T(t_k)] = Q$$

在式(1)中, 由于每部雷达的测量误差不尽相同, 因此 $\sigma_w^2(i) (i = 1, 2, \dots, N)$ 可能两两不等, 作线性组合:

$$Z(t_k) = \sum_{i=1}^N \alpha_i Z^i(t_k) \quad (2)$$

且令 $\sum_{i=1}^N \alpha_i = 1$, 可以得到^[5]:

$$\alpha_i = \frac{1}{\sigma_w^2(i) \sum_{j=1}^N \frac{1}{\sigma_w^2(j)}} \quad (3)$$

2.2 目标跟踪

基于式(1)对目标进行跟踪, 进行目标状态估计和预测。卡尔曼滤波是一种线性、无偏、以误差方差最小为准则的最优估计算法, 利用卡尔曼滤波器

可得到基于融合测量的目标状态滤波值。

滤波估值方程为:

$$\hat{X}(k+1|k+1) = \hat{X}(k+1|k) + K(k+1) [Z(k+1) - H(k+1)\Phi(T, \alpha)\hat{X}(k|k)] \quad (4)$$

预测估值方程为:

$$\hat{X}(k+1|k) = \Phi(T, \alpha)\hat{X}(k|k) \quad (5)$$

增益方程为:

$$K(k+1) = P(k+1|k)H^T(k+1) [H(k+1) \times P(k+1|k)H^T(k+1) + R(k+1)]^{-1} \quad (6)$$

滤波估值协方差矩阵方程为:

$$P(k+1|k+1) = [I - K(k+1)H(k+1)]P(k+1|k) \quad (7)$$

预测估值协方差矩阵方程为:

$$P(k+1|k) = \Phi(T, \alpha)P(k|k)\Phi^T(T, \alpha) + Q(k) \quad (8)$$

3 仿真结果

应用式(4)~(8)的目标跟踪算法, 取观测转移矩阵为:

$$\Phi(T, \alpha) = \begin{bmatrix} 1 & T & \frac{T^2}{2} & 0 & 0 & 0 \\ 0 & 1 & T & 0 & 0 & 0 \\ 0 & 0 & e^{-\alpha T} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & T & \frac{T^2}{2} \\ 0 & 0 & 0 & 1 & T & 0 \\ 0 & 0 & 0 & 0 & 0 & e^{-\alpha T} \end{bmatrix} \quad (9)$$

取观测噪声协方差矩阵为:

$$Q(k) = \begin{bmatrix} q_{11} & q_{12} & q_{13} & 0 & 0 & 0 \\ q_{21} & q_{22} & q_{23} & 0 & 0 & 0 \\ q_{31} & q_{32} & q_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & q_{11} & q_{12} & q_{13} \\ 0 & 0 & 0 & q_{21} & q_{22} & q_{23} \\ 0 & 0 & 0 & q_{31} & q_{32} & q_{33} \end{bmatrix}$$

式中

$$q_{11} = (1 - e^{-2\alpha T} + 2\alpha T + 2\alpha^3 T^3 / 3 - 2\alpha^2 T^2 - 4\alpha T e^{-\alpha T}) / (2\alpha^5)$$

$$q_{12} = (1 + e^{-2\alpha T} - 2e^{-\alpha T} + 2\alpha T e^{-\alpha T} - 2\alpha T + \alpha^2 T^2) / (2\alpha^4) = q_{21}$$

$$q_{13} = (1 - e^{-2\alpha T} - 2\alpha T e^{-\alpha T}) / (2\alpha^3) = q_{31}$$

$$q_{22} = (4e^{-\alpha T} - 3 - e^{-2\alpha T} + 2\alpha T) / (2\alpha^3)$$

$$q_{23} = (1 + e^{-2\alpha T} - 2e^{-\alpha T}) / (2\alpha^2) = q_{32}$$

$$q_{33} = (1 - e^{-2\alpha T}) / (2\alpha)$$

取估值协方差矩阵初值为:

$$P_0 = \begin{bmatrix} \sigma_{x2}^2 & \frac{\sigma_{x2}^2}{T} & 0 & \sigma_{xy2}^2 & \frac{\sigma_{xy2}^2}{T} & 0 \\ \frac{\sigma_{x2}^2}{T} & \frac{\sigma_{x1}^2 + \sigma_{x2}^2}{T^2} & 0 & \frac{\sigma_{xy2}^2}{T} & \frac{\sigma_{xy1}^2 + \sigma_{xy2}^2}{T^2} & 0 \\ 0 & 0 & \sigma_m^2 & 0 & 0 & 0 \\ \sigma_{xy2}^2 & \frac{\sigma_{xy2}^2}{T} & 0 & \sigma_{y2}^2 & \frac{\sigma_{y2}^2}{T} & 0 \\ \frac{\sigma_{xy2}^2}{T} & \frac{\sigma_{xy1}^2 + \sigma_{xy2}^2}{T^2} & 0 & \frac{\sigma_{y2}^2}{T} & \frac{\sigma_{y1}^2 + \sigma_{y2}^2}{T^2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_m^2 \end{bmatrix}_{6 \times 6}$$

取量测矩阵为：

$$H(k) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

应用1.3节的网格聚类算法，以及上述的目标跟踪算法及雷达观测数据融合方法，以三部Alenia雷达录取的数据截取一段进行处理，得到的融合航迹如图2所示。

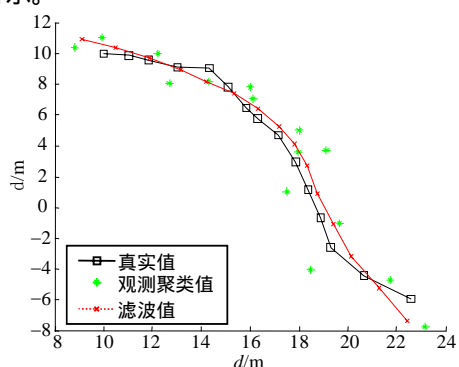


图2 网格聚类滤波跟踪航迹

4 结束语

本文讨论了网格聚类的基本思想、聚类处理过程及具体的算法实现步骤，在对观测数据聚类处理基础上，讨论了数据融合和目标跟踪的算法，确定了滤波方程中的观测转移矩阵、观测噪声协方差矩阵、估值协方差矩阵初值及量测矩阵。通过MATLAB仿真，得到了聚类处理后的滤波跟踪航迹图，仿真

结果表明，在观测值聚类处理后的滤波跟踪航迹效果较好，网格聚类在处理观测值噪声方面有较好的作用。

参考文献

- [1] MACQUEEN J. Some methods For classification and analysis of multivariate observations[C]// Proc 5th Berkeley Symp Math Statist Prox. Berkeley: University of California, 1967: 281-297.
- [2] KAUFMAN L, ROUSSEUW P J. Finding groups in data an introduction to cluster analysis[C]//IEEE Transactions on Pattern Analysis and Machine Intelligence. New York: John Wiley&Sons, 1990: 3: 134-148.
- [3] 王海, 王忠民. 一种基于密度和网格的聚类算法在KDD中的应用[J]. 计算机工程与应用, 2004, 24(3): 179-183.
- [4] 田启明, 王丽珍, 尹群. 基于网格距离的聚类算法的设计、实现和应用[J]. 计算机应用, 2005, 25(2): 295-297.
- [5] 赵国庆. 雷达对抗原理[M]. 西安: 西安电子科技大学出版社, 1999.
- [6] 舒红平. 空中交通管理系统中航迹发散判据及抑制处理[J]. 控制与决策, 2005, 20(3): 345-348.
- [7] 刘洋, 徐毓. 基于观测值聚类多雷达数据融合[J]. 雷达与对抗, 2002, 2(1): 5-11.
- [8] SHU Hong-ping, XU Zheng-ming, SOU Shu-rong, et al. Research on Multiradar Data Fusion Algorithm based on K-central Clustering[C]//Proceedings of 2006 Symposium on Information(IECT'2006). Chengdu: Southwest Jiaotong University Press, 2006: 47-50.

编辑 税红