

基于自动机模型数据关联性能评估算法

梁冰¹, 刘群²

(1. 哈尔滨理工大学计算机科学与技术学院 哈尔滨 150001; 2. 哈尔滨工程大学计算机科学与技术学院 哈尔滨 150001)

【摘要】提出了一种基于自动机理论的性能评估方法,通过在有限自动机模型中增加描述观测的时间序列,来扩展其对数据关联过程的表达能力;使用时序有限自动机模型来描述关联处理算法的模型;使用带时间属性的字符串表示目标航迹。对数据关联的时序有限自动机模型进行一致性和可达性分析,并在此基础上给出了性能评估算法。实验结果证明该方法可以有效地评价关联算法的性能。

关键词 数据关联; 性能评估; 时序有限自动机模型; 时间序列
中图分类号 TP213 **文献标识码**

Method of Data Association Performance Evaluation Based on Automata Theory Consistency and Reachability Analysis

LIANG Bing¹, and LIU Qun²

(1. College of Computer Science and Technology, Harbin University of Science and Technology Harbin 150001;
2. College of Computer Science and Technology, Harbin Engineering University Harbin 150001)

Abstract The research goal of data association in the past has primarily focused on establishing a computational approach for association processing algorithms. In this paper, a performance evaluation model is proposed for data association based on temporal finite automata. An algorithm for checking the consistency and reachability of the model is developed. A simulation experiment illustrates the effectiveness of this method in the analysis and verification of the performance of data association algorithms.

Key words data association; performance evaluation; temporal finite automata; time series

在过去20多年里,多目标跟踪在理论和应用两个方面都取得了长足的进展。数据关联作为目标跟踪的核心部分,目前的工作更多地集中在各种关联算法方面的研究,如最近邻域法^[1]、联合概率互联算法^[2]、多假设算法^[3]等,而关联算法的性能评估方面则较少得到关注。由于关联的准确性是影响目标跟踪系统性能的重要因素^[4-7],如何借助自动机理论分析和评价不同的关联计算,是本文主要研究的问题。

有限自动机是描述算法计算模型的有力工具,根据不同的应用得到了很多扩展研究,有效地支持了各种复杂行为的建模^[8-11]。鉴于一个数据关联过程中的输入是按照确定的采样周期,这就存在一个时序要求,本文定义了语意更为丰富的时序有限自动机模型来描述。建立起关联算法的时序有限自动机模型后,通过对数据关联时序有限自动机模型的一致性和可达性分析,检验跟踪系统所输出航迹的精确度和完整性,使之成为评估数据关联算法的数学

工具。

1 数据关联问题的描述

数据关联可以看作是建立来自不同传感器、不同时刻的观测与其他观测数据的关系,以确定它们是否来自同一个数据源(即目标或事件)的处理过程。对数据关联问题进行形式化的描述,主要有两个方面,一是对来自不同时刻的观测数据进行符号定义;二是对航迹以字符串形式定义。

传感器周期性的接收到若干观测,假设在 $t_0=0$ 时刻开始扫描,那么在 $[t_0, t_1]$ 内获得第一个观测值的集合,其中 t_1 是第二次扫描开始的时间,假设每次扫描的时间间隔 T 是相等的。第 k 次扫描结束的时间 $t_k=kT$ 。但周期长度 T 固定时,时间可只考虑为周期的个数,记为 t ,则第 t 次扫描得到的一组观测值表示为:

$$C(t) = \{z_t^k\} \quad k=1, 2, \dots, g(t) \quad (1)$$

式中 z 为观测值; t 为周期个数或时间长度; k 为本

收稿日期: 2006-11-06; 修回日期: 2007-06-18

作者简介: 梁冰(1981-), 女, 博士, 主要从事多传感器融合方面的研究; 刘群(1957-2007), 男, 教授, 博士生导师, 主要从事人工智能、多传感器融合方面的研究。

次扫描的观测在组中的序号; $g(t)$ 是 t 次扫描获得的观测值个数, 不同周期可能获得观测值数量不同, 故为 t 的函数。

数据关联的实质是把传感器在不同时刻扫描到的观测数据进行分类, 在每一类按时序可以形成轨迹, 分别对应着一个运动的目标。对数据关联问题的形式化描述为:

若干次扫描构成了一个观测数据集的时间序列, 记作 $Z(t) = \{C(1), C(2), \dots, C(t)\}$ 。 $Z(t) = \bigcup_{i=1}^t C(i)$, 表示成:

$$Z(t) = \{z_i^k\}^t \quad (2)$$

关联的处理是对 $Z(t)$ 中的观测形成正确的分类, 把来自扫描的每组观测值 $C(i)(i=1, 2, \dots, t)$ 中表示同一个目标的观测值归并到一起, 使它们构成描述这个目标的观测值集合。每一类记作 $D(i)^t = \{d|d \text{ 来自目标 } i \text{ 的观测}\}(i=1, 2, \dots, n)$ 。 $t+1$ 时刻, 再对 $C(t+1)$ 中的元素进行分类, 使得 $D(i)^t = D(i)^t \cup \{z_{t+1}^{i(t+1)}\}$ 。航迹 i 的时序字符串的形成是通过对 $D(i)^t$ 中来自不同时刻观测的处理, 按照时间顺序依次排列形成的。如 $D(i)^n = \{z_1^{i(1)}, z_1^{i(2)}, \dots, z_n^{i(n)}\}$, 航迹 i 表示为 $z_1^{i(1)} z_2^{i(2)} \dots z_n^{i(n)}$, $i(1), i(2), \dots, i(n)$ 表明同一目标的观测值在不同扫描周期中的观测序号是不同的。一个数据关联算法得到 n 条航迹, 就对应 n 个字符串。

2 数据关联模型

2.1 时序有限自动机

传感器对观测区域的周期性采样获得了目标的时序观测值, 数据关联处理的就是观测数据具有时间的属性, 因此为了扩展有限自动机的功能, 定义一个时间序列来描述传感器的采样周期并作为时序有限自动机一项。

定义 1 时间序列 $T = \{t_1, t_2, \dots, t_n\}$, 其中 $t_1 < t_2 < \dots < t_n$, $t_n \in N$ (N 是自然数集)。

定义 2 时序有限自动机 (temporal finite automata, TFA) 是一个六元组 $M = (T, Q, \Sigma, \delta, (q_0, 0), F)$, 其中:

(1) 一个时间序列 $T = \{t_1, t_2, \dots, t_n\}$;

(2) 时序状态集 $Q \times T = \{(q_i, t_j) : q_i \in Q, t_j \in T\}$, q 是有限状态集合, TFA的状态由有序对 (q_i, t_j) 标记, 它的第一个元素是状态, 第二个元素是时刻;

(3) $\Sigma \times T = \{(z_i, t_j) : z_i \in \Sigma, t_j \in T\}$, Σ 为输入字母表, $\Sigma \times T$ 表示TFA在时刻 t 输入字母表, TIOA的输入由有序对 (z_i, t_j) 标记;

(4) $\delta : (Q \times T) \times (\Sigma \times T) \rightarrow (Q \times T)$ 的有穷子集, 称为状

态转移函数, 可表示为:

$$\begin{aligned} \delta((q_i, 0), (z_i, t')) &= (q_i, t'') \\ \delta((q_i, t), (z_i, t')) &= (q_i, t'') \end{aligned}$$

在时序性上, 满足时间函数:

$$f(t) = \begin{cases} t' = t + c \\ t'' = t' \end{cases}$$

式中 $t, t', t'' \in T$, c 为非负整数;

(5) $(q_0, 0)$ 为初始状态;

(6) F 为终止状态集合。

2.2 数据关联的TFA模型

在单目标情况下, 数据关联可以表示成一个基本的时序有限自动机, 它由一个传感器观测的时间序列 T 、关联处理的初始状态 $(q_0, 0)$ 、目标的时序观测数据集 $\Sigma \times T$, 按照观测与航迹的配对关系设定状态转移函数 δ , 关联处理的终止状态 (q_f, n) 组成。单目标的TFA模型可表示为:

$$\begin{aligned} M &= (T, Q, \Sigma, \delta, (q_0, 0), (q_f, n)) \\ T &= \{1, 2, \dots, n\} \\ Q \times T &= \{(q_i, t_j) : q_i \in Q, t_j \in T\} \\ \Sigma \times T &= \{(z_i, t_j) : z_i \in \Sigma, t_j \in T\} \end{aligned}$$

状态转移函数 $\delta((q_0, 0), (z, t)) = (q_i, t)$ 表示在 t 时刻, 观测 z 关联到目标的航迹, 如在 t 时刻, 没有观测关联到目标航迹, 状态转移函数为 $\delta((q, 0), (\varepsilon, t)) = (q, t)$ 。如果航迹在连续5次扫描中都没有关联到非零的观测, 则认为航迹终止。终止状态 (q_f, t) 表示航迹结束, t 表示航迹的终止时刻。时序有限自动机的一条专一路径如下: $(q_0, 0) \xrightarrow{(z_1, t_1)} (q_1, t_1) \rightarrow \dots \xrightarrow{(\varepsilon, t_i)} (q_i, t_i) \dots \xrightarrow{(z_n, t_n)} (q_f, t_{\text{end}})$, 该路径表示单目标的一条航迹 $(z_1, t_1) \dots (\varepsilon, t_i) \dots (z_n, t_n)$, 航迹在 t_i 时刻没有观测关联, 在 t_n 时刻终止。TFA接受的语言是目标的一条航迹。扫描的次数决定了航迹的长度, 按时间顺序排列真实目标的观测, 从而得到目标航迹的延续。

在多目标情况下, 数据关联的TFA模型有一个初始状态和多个终止状态, 初始状态表示关联处理的开始, 而多个结束状态表示多个目标航迹的结束。模型中的多条路径表示对不同目标观测数据关联处理过程, 而接受的语言集合包含多个目标的航迹。多目标情况下, 在某一时刻, 关联算法的约束条件为: (1) 一个观测至多只能属于一个目标; (2) 一个目标至多只能拥有一个量测。因此, 状态转移函数满足如下条件: $\delta((q_i, t), (z, t')) = (q_i, t')$ 不能在同一时刻出现, 保证一个观测至多只能输入同一目标; $\delta((q_i, t), (z, t')) = (q_i, t')$ 和 $\delta((q_i, t), (c, t')) = (q_i, t')$ 不能在同一时刻出现, 保证一个目标至多只能拥有一

个观测。

3 模型分析

建立数据关联过程的时序有限自动机模型后,可以通过对模型的分析,来评估关联算法的性能。如通过检查模型的一致性和可达性,确定关联算法的完整性和正确性。

3.1 一致性分析

一致性验证问题,即某个真实的目标对应的航迹在关联算法输出的航迹中至少出现一次。由于传感器的漏检、杂波的干扰、关联的方法等各种原因,造成真实目标没有对应的系统输出航迹,目标与航迹之间存在不一致现象,使用目标跟踪系统性能不能达到要求。其中,最主要的就是航迹的遗漏,也就是说,需要检查在TFA描述的所有航迹中,真实目标的航迹集合描述的语言 $L(T)$ 是否至少出现一次。

3.2 可达性分析

可达性问题,即关联算法得到了真实目标的正确航迹。也就是检验一条真实目标轨迹 $\omega \in L(T)$ 是否在自动机中存在一条标记为 $x \in \Sigma^*$ 的路径,使得TFA从初始状态到达终止状态。可达性描述了系统输出航迹的正确性。正确性是指关联算法能正确地将观测和目标进行配对是跟踪系统输出航迹的一个重要特征。一个错误的关联,可能会严重地损害目标航迹和以后的状态估计的准确性,因此关联算法的正确性影响到系统的优劣。构造关联算法的时序有限自动机后,正确性实际上是其对应的TFA在接收到表示为自动串的真实目标轨迹后,从其唯一的初始状态到终止状态的可达性。

定义 3 数据关联的时序有限自动机模型可达性是指若存在一条航迹,可将时序有限自动机的状态从 $(q_0, 0)$ 转换为 (q_n, t) 时,则状态 (q_n, t) 称为从状态 $(q_0, 0)$ 可达。

3.3 性能评估算法

基于以上分析,本文给出相应的性能评估算法。TFA描述的是关联算法得到所有可能的航迹(即输入观测序列)。在此算法中从初始节点对状态转移图进行遍历,将访问过的航迹,根据访问的结果放在不同的集合中, missed_track[] 集合中记录了某些遗漏的航迹; correct_track[] 集合中记录了某些正确的航迹, false_track[] 集合中记录了某些错误的航迹。记录的出错信息可以用来对跟踪系统进一步理解和对关联算法的改进。

一致性和可达性分析算法都是基于自动机的状

态转移图进行广度优先或深度优先搜索过程,其顶点为符号化的状态。在状态空间的搜索过程中存在宽度搜索和深度搜索两种方法。从初始状态开始,在展开状态空间的同时,检测新生成的状态是否满足性质,若满足,则搜索过程结束;否则,继续对新生成的状态展开所有的后继,直到不再生成新的状态为止,整个过程结束。下面根据搜索的方式分析数据关联的TFA模型的算法。

输入: 关联算法的时序有限自动机模型, 目标真实航迹集合 TRACK {w_{id} | id为目标标实};

输出: 正确航迹, 遗漏航迹, 错误航迹集合。

```
Present.Add([q0,0]);
Num= GetTrackSequenceSize();
for(int Index=0; Index<Num; Index++)
    for(int P=0; P<Present.GetSize(); P++)
        TransConnState(ptrConnNode,w); //按照关联
        处理的状态转换函数对连接进行状态迁移
        if(Present.HasstartState())
            then将TRACK wid放入missed_track//此航
            迹没有使得TFA发生状态改变,表明关联算法遗漏
            此目标的航迹;
            if(!HasExist(Trans,Trsifion))
                GetConditionValue(Trsition)
                Trans.Add(Trsition)
            End if
        End for//生成与Present相连的转移Trans
        for(int T=0;T<Trans.GetSize();T++)
            Loca=Trans.GetAt(T).GetConnLoca();
            if(!HasExist(Next, Loca))
                Next.Add(Loca);
            Endif
        End for//生成与转移Trans相连的状态Next
        if(Next.IsEmpty())
            then 将TRACK wid放入false_track//没有路径可
            走,表明关联算法得到的目标航迹有错误,即关联不
            属于本目标的观测;
            Endif
        Present.RemoveAll();
        Trans.RemoveJAll();
        Present=Next;
        End for
        if(Present.HasFinalState())
            将TRACK wid放入correct_track//此航迹使得
            TFA到达了终止状态,航迹被接受,表明关联算法得
```

到了目标id的一条正确航迹;

else

将TRACK w_{id} 放入false_track;

End if

算法对关联算法的到达航迹进行分析后, 通过对missed_track[]、correct_track[]、false_track[]集合中遗漏航迹、正确航迹、错误航迹的统计, 得出目标跟踪系统正确性和完整性的指标, 正确航迹概率用正确航迹数与航迹总数比值来描述, 遗漏航迹概率用遗漏航迹数与目标总数比值来描述。

4 仿真实验

实验的设计为: 模拟的传感器是一个单扫描监视雷达, 在50 km范围内提供二维的位置数据。目标数分别为10批、30批、50批、80批、100批, 扫描每个场景中的目标为30次。雷达采样时间间隔为8 m。每周杂波个数按泊松分布 $\lambda_p=5$ 。观测噪音在横坐标和纵坐标方向, 分别设为 $\sigma_r=0.015$ km和 $\sigma_b=0.005$ 2 km。此外, 目标观测概率为0.90。目标做匀速直线运动, 速度随机在0.1~1 km/s之间, 目标的初始位置均匀地分布在100 km×100 km的区域内。分别用最近邻域法(NN)、概率数据关联算法(PDA)和联合概率互联算法(JPDA)对场景中数据进行关联, 得到的结果用评估算法进行评估。

实验在P4、CPU 3.00 GHz、512 MB内存的Win2000系统下进行, 航迹的长短和目标的数量是影响算法时间消耗的关键因素。表1为算法在不同目标数下的运行时间值。由表1的结果可以看出, 随着目标个数的指数递增, 运行的速度没有呈指数上升, 而是线性上升。

表1 算法在不同目标数下的运行时间

目标数	运行时间/s		
	NN	PDA	JPDA
10	0.52	0.60	0.67
30	1.62	1.84	2.05
50	2.56	2.83	3.14
80	4.12	4.23	4.40
100	5.25	5.36	5.52

图1为正确航迹概率图。从图1可知3种算法得到的航迹都比实际目标数要大, 这是因为存在一些虚假目标或是冗余的航迹, 随着目标密度的增大, 3种

算法得到的正确航迹都有所下降, 但算法JPDA仍取得较好的效果。

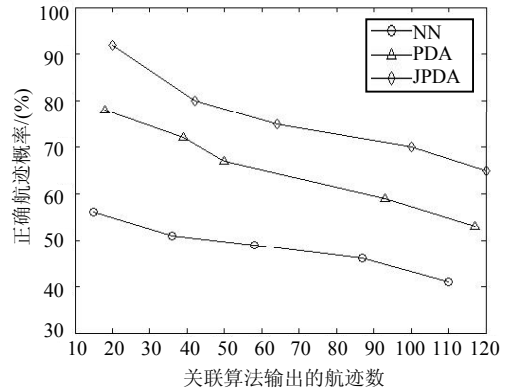


图1 正确航迹概率

图2为遗漏航迹概率图。从图中可知, NN在目标密度大的情况下遗漏了较多的航迹, JPDA也有部分遗漏航迹, 但在可接受的范围。从各个关联算法的性能分析图可以看出3个算法的性能特点: NN在密集目标或杂波干扰环境下, 其错误较多; PDA、JPDA在杂波环境中对多目标跟踪是相对理想的方法。

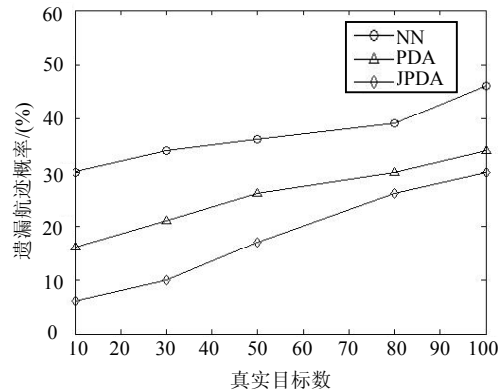


图2 遗漏航迹概率

5 结论

本文提出一种时序有限自动机模型, 来描述目标跟踪系统中的数据关联, 给出了模型的一致性和可达性分析方法。基于该模型所实现的评估算法能够初步地进行数据关联算法的评估, 算法在时间和空间上随着目标的增长而呈线性增长; 在性能分析方面, 得出了相应的性能评价指标, 反映出关联算法的正确性和完整性。下一步的工作是优化数据关联的时序有限自动机模型, 寻求更有效的模型分析方法。

(下转第629页)

明GFCMA算法要优于FCMA算法。

表1 3幅图片的FCMA和GFCMA算法计算结果

图片	类别数	FCMA均方误差	GFCMA均方误差
a	3	12.287 6	12.238 5
b	3	11.816 1	10.472 4
g	3	9.652 3	10.171 8

4 结 束 语

从本文的算法过程可以看出, 当需要处理的问题的样本或样本维数很大时, 适应度函数的计算量变得非常巨大, 导致算法需要花费大量的时间才能收敛到最优值。为减少计算量, 本文中实际计算时, 将遗传算法与模糊C-均值聚类方法结合起来。具体就是初始化群体后, 先对群体中的每个个体应用FCMA迫使其收敛到局部最优, 然后再进行交叉变异操作; 最后再对遗传算法中每一代交叉变异所产生的新个体在可能被加入群体之前也先应用FCMA使之收敛到局部最优。然而尽管作如此的改进, GFCMA算法仍是相当耗时的一项工作。

参 考 文 献

[1] VELTHUIZEN R P, HALL L O, CLARKE L P. Feature extraction for MRI segmentation[J]. Neuro Imaging, 1999, 9(2): 85-90.
 [2] BARRA V, BOIRE J V. Tissue segmentation on MR images of the brain by possibilistic clustering on a 3D wavelet representation[J]. Magn Reson Imag, 2000, (11): 267-278.

[3] SCHROETER P, VESIN J M, LANGENBERGER T, et al. Robust parameter estimation of intensity distribution for brain magnetic resonance images[J]. IEEE Trans Med Imaging, 1998, 17(2): 172-186.
 [4] PHAM D L, PRINCE J L. Adaptive fuzzy segmentation of magnetic resonance images[J]. IEEE Trans Med Imaging, 1999, 18: 737-752.
 [5] 章 晋. 图像分割[M]. 北京: 科学出版社, 2001. ZHANG Jin. Images Segmentation[M]. Beijing: Science Press, 2001.
 [6] 雷英杰, 张善文, 李续武, 等. 遗传算法工具箱及应用[M]. 西安: 西安电子科技大学出版社, 2005. LEI Ying-jie, ZHANG Shan-wen, LI Xu-wu, et al. Genetic algorithm tools & applications[M]. Xi'an: Xidian University Press, 2005.
 [7] WANSCHURE T, COLEY D A, VENNART W, et al. Automat realignment of time-separated MR images by genetic algorithm[J]. Magnersson Imaging, 1999, 17(2): 313-317.
 [8] 周叙国, 邢汉承. 基于遗传模糊C-均值聚类算法的地图分割[J]. 计算机与数字工程, 2005, 33(6): 77-78. ZHOU Xu-guo, XING Han-cheng. Topographic map segmentation based on genetic fuzzy C-means clustering algorithm[J]. Computer & Digital Engineering, 2005, 33(6): 77-78.
 [9] 边肇祺, 张学工. 模式识别[M]. 北京: 清华大学出版社, 2004. BIAN Zhao-qi, ZHANG Xue-gong. Pattern recognition[M]. Beijing: Tsinghua University Press, 2004.

编 辑 熊思亮

(上接第609页)

参 考 文 献

[1] DEMING R, PERLOVSKY L. Concurrent multi-target localization, data association, and navigation for a swarm of flying sensors[J]. Information Fusion, 2007, 8(3): 316-330.
 [2] ZHU H. The algorithm and simulations for the asynchronous track-to-track association[C]//Proceedings of the Sixth International Conference on Information Fusion. Washington: IEEE, 2003: 1400-1404.
 [3] OUSSALAH M, SCHUTTER J. Hybrid fuzzy probabilistic data association filter and joint probabilistic data association filter[J]. Information Sciences, 2002, 142(1): 195-226.
 [4] MEGALOOIKONOMOU V, YESHA Y. Space efficient quantization for distributed estimation by a multi-sensor fusion system[J]. Information Fusion, 2004, 5(4): 299-308.
 [5] DRUMMOND E. Performance Metrics for Multiple-Sensor, Multiple -Target Tracking[C]//Signal and Data Processing of Small Targets 2000, Proceedings of SPIE. Atlanta: IEEE, 2000.
 [6] CHANG K, TIAN Z MAP track fusion performance evaluation[C]//Proc Fusion 2002. Washington DC: [s.n.],

2002.
 [7] REID D. Optimising the flow of information within a C3I network[J]. Mathematical and Computer Modelling, 1999, 30(9): 225-243.
 [8] SONG H, Compton K. Rounds SPIN: a model checker for reconfigurable hybrid systems based on SPIN[J]. Electronic Notes in Theoretical Computer Science, 2006, 145(14): 167-183.
 [9] VENEMA Y. Automata and fixed point logic: a coalgebraic perspective[J]. Information and Computation, 2006, 204(4): 637 -678.
 [10] 高 军, 杨冬青. 基于树自动机的XPath在XML数据流上的高效执行[J]. 软件学报, 2005, 16(2): 223-232. GAO Jun, YANG Dong-qing. Tree automata based efficient XPath evaluation over XML data stream[J]. Journal of Software, 2005, 16(2): 223-232.
 [11] DIERKS H. PLC-automata: a new class of implemenTab real-time automata[J]. Theoretical Computer Science, 2001, 253: 61-93.

编 辑 税 红