

文章编号: 1006 - 4729(2010)01 - 0079 - 04

基于 MHT 的多物体跟踪算法

邵 洁

(上海电力学院 计算机与信息工程学院, 上海 200090)

摘 要: 探讨了多假设跟踪算法(MHT)在多物体检测跟踪系统中的应用方式. 该方法采用卡尔曼滤波算法实现跟踪轨迹假设, 采用 LAP 算法实现数据关联, 将预测和视频采集测量值结合得到最优跟踪轨迹. 实验表明, 该方法在实践中达到了较好的效果.

关键词: 卡尔曼滤波; 多假设跟踪; 线形分配; 多物体跟踪
中图分类号: TP391.41 **文献标识码:** A

Multiple Target Tracking Based on MHT

SHAO Jie

(School of Computer and Information Engineer, Shanghai University of Electric Power, Shanghai 200090, China)

Abstract: An application of Multiple hypothesis tracking(MHT) in multiple target tracking system is discussed. Algorithm of Kalman filtering is used to get tracking hypothesis and LAP algorithm is used to implement data association. Optimal tracking results are obtained by the combination of the hypothesis and the measurement of image sequences. Experiments show that the method is effective in practice.

Key words: Kalman filter; MHT; LAP; multiple object tracking

运动目标分析是计算机视觉研究领域中的一个富有挑战性的课题. 它涵盖了移动目标物体的检测、目标物体的跟踪, 以及跟踪物体的行为识别等方面的内容. 这一课题在医学检测、军事监测、安全报警、智能家居、录像检索与回放等高科技领域具有广泛的应用前景.

1 多物体跟踪系统

目标跟踪是指将收集在一定区域内的一个或多个潜在目标的传感信息进行判断, 区分出新目标和匹配原始跟踪路径的已有目标. 更重要的是, 一旦跟踪确立, 就可以进一步估计出目标物体的

大量有用信息, 例如运动速度、下一帧行进位置、目标分类特征等. 一般而言, 常见的监控系统都是针对多物体场景设置的, 因此多物体跟踪(MTT)是最重要的跟踪应用模式. 图 1 为典型的 MTT 系统的基本流程图^[1].

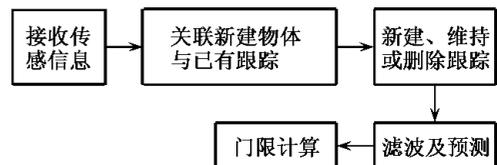


图 1 MTT 系统基本流程

收稿日期: 2009 - 09 - 02

作者简介: 邵洁(1981 -), 女, 硕士, 讲师, 江苏南京人. 主要研究方向为计算机视觉. E-mail: shaoj. shiep@gmail.com.

当前,多目标跟踪的难点主要来源于以下几个方面:一是突发性物体运动;二是运动目标和运动场景同时变化;三是目标结构运动过程中的形变;四是目标间或目标与背景间产生遮挡;五是摄像机移动状态下的运动跟踪.

根据 ALPER Yilmaz 等人^[2]的统计,能够实现多物体跟踪算法的主要有来自于 SETHI 和 JAIN 的最近邻法(GNN)、SHAFIQUE 的多帧数据跟踪法(MFT)、BAR-Shalom 的联合概率数据关联算法(JPDAF)等.

与这些方法相比,多假设跟踪算法(MHT)同时具有实现多目标场景下的物体进入、离开,以及被短暂遮挡等特殊情况下的跟踪能力,同时易于实现程序优化.目前,国外许多多物体跟踪系统都广泛采用 MHT 作为核心算法.

1978年, DONALD B Reid 最先提出了用于跟踪红外成像目标的多假设跟踪算法(MHT). MHT 是一种在数据关联发生冲突时形成多种假设以延迟做决定的逻辑.

但随着假设数量的增加,计算量将成指数增长,因此本文选择优化后的基于轨迹的多假设跟踪算法^[3],将其流程分为两个步骤——数据关联和航迹维护.实现步骤如图2所示.

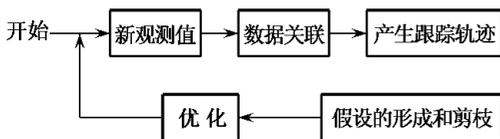


图2 MHT基本实现流程

在实际应用中,本文采用卡尔曼滤波算法^[4]实现跟踪轨迹假设,采用 LAP(Linear Assignment Problem)算法实现数据关联.

2 卡尔曼滤波算法

一个离散控制过程的系统可用一个线性随机微分方程来描述:

$$X(k) = A X(k-1) + B U(k) + W(k) \quad (1)$$

同样,对于系统在某一时刻的测量采样值也可以用线性随机微分方程来描述:

$$Z(k) = H X(k) + V(k) \quad (2)$$

式中: $X(k)$ —— k 时刻的系统状态;

$U(k)$ —— k 时刻对系统的控制量;

A, B —— 系统参数,对于多模型系统,为矩

阵;

$Z(k)$ —— k 时刻的测量值;

H —— 测量系统的参数,对于多测量系统,为矩阵;

$W(k), V(k)$ —— 系统噪声和测量噪声,一般可以被假设成高斯白噪声,其协方差矩阵分别为 Q 和 R (这里假设其不随系统状态变化而变化).

对于满足以上条件的线性随机系统,卡尔曼滤波器是最优的信息处理器.利用卡尔曼滤波器,结合测量值和系统估计值的协方差,可以估算出系统的最优化输出.

假设当前状态为 $x(k)$,则可根据上一状态的最优预测 $X(k-1|k-1)$,给出当前状态的预测值:

$$X(k|k-1) = A X(k-1|k-1) + B U(k) \quad (3)$$

更新相对应于预测值 $X(k|k-1)$ 的协方差:

$$P(k|k-1) = A P(k-1|k-1) A' + Q \quad (4)$$

结合式(2),计算当前状态下的最优预测值:

$$X(k|k) = X(k|k-1) + K_g(k) [Z(k) - H X(k|k-1)] \quad (5)$$

计算当前状态的卡尔曼增益:

$$K_g(k) = \frac{P(k|k-1) H'}{H P(k|k-1) H' + R} \quad (6)$$

最后更新当前状态下最优预测值 $X(k|k)$ 的协方差:

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

3 基于 MHT 的多物体跟踪算法的实现

MHT 实现多假设跟踪的方式示例 1 如图 3 所示.

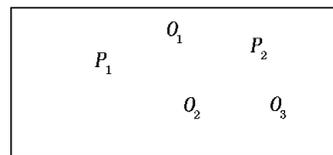


图3 视频跟踪场景示例1

假设当前视频第 $k-1$ 帧已有 2 个跟踪目标,分别为 T_1, T_2 ,图3中的 P_1, P_2 分别为 T_1, T_2 在第 k 帧根据式(3)和式(4)求得的假设位置. $O_1, O_2,$

O_3 为 k 帧观测到的 3 个物体,这里提取观测物体可以采用多种背景提取方法,如帧差法、高斯混合模型(GMMs)等,本文采用 GMMs 提取前景图像.根据当前情况,可能出现的匹配假设有 10 个^[5],分别为 $H_1 - H_{10}$. 定义 $T_n(P_1, O_1)$ 表示跟踪轨迹 T_n 来自于 P_1 和 O_1 应用式(5)至式(7)关联得到的最优估计. $T_n(O_3)$ 表示根据观察新产生的跟踪路径 T_3 . $T_n(P_1)$ 表示无观察物体与之匹配,以假设值延续跟踪路径 T_n .

$$\begin{aligned} H_1: & T_1(P_1, O_1); T_2(P_2, O_2); T_3(O_3). \\ H_2: & T_1(P_1, O_2); T_2(P_2, O_1); T_3(O_3). \\ H_3: & T_1(P_1); T_2(P_2); T_3(O_1); T_4(O_2); \\ & T_5(O_3). \\ & \dots \end{aligned} \quad (8)$$

3.1 数据关联

假设视频第 $k-1$ 帧图像已有跟踪轨迹数目为 N_{k-1} ,第 k 帧观测到物体数目为 m_k ,则其相关性矩阵 $\Omega(H^k)$ 由式(9)确定:

$$\Omega(H^k) = \begin{bmatrix} \omega_{1,1}^k & \omega_{1,2}^k & \dots & \omega_{1,j}^k & \dots & \omega_{1,N_{k-1}}^k \\ \omega_{2,1}^k & \omega_{2,2}^k & \dots & \omega_{2,j}^k & \dots & \omega_{2,N_{k-1}}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ \omega_{i,1}^k & \omega_{i,2}^k & \dots & \omega_{i,j}^k & \dots & \omega_{i,N_{k-1}}^k \\ \vdots & \vdots & & \vdots & & \vdots \\ \omega_{m_k,1}^k & \omega_{m_k,2}^k & \dots & \omega_{m_k,j}^k & \dots & \omega_{m_k,N_{k-1}}^k \end{bmatrix} \quad (9)$$

式中: j ——行标号,对应当前帧的测量向量 z_{kj} ;
 i ——列标号,对应已有跟踪轨迹 T_i ;
 ω_{ij}^k ——用于判断测量值和轨迹的相关性,根据实际应用,以轨迹中心点和测量物中心点间欧式距离定义 ω_{ij}^k 的值.

采用 LAP 算法化简矩阵 $\Omega(H^k)$,得到 $\Omega'(H^k)$ 实现:

$$\omega_{ij}^k = \begin{cases} 1 & z_{kj} \in T_i \\ 0 & z_{kj} \notin T_i \end{cases} \quad (10)$$

按图 3 所示情况,可以得到:

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

3.2 跟踪轨迹的形成与修剪

相关性化简矩阵 $\Omega'(H^k)$ 的结果在某些情况

下需要修正. MHT 实现多假设跟踪的方式示例 2 如图 4 所示.

按图 4 所示情况将会得到:

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

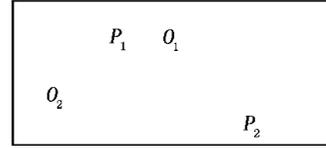


图 4 视频跟踪场景示例 2

事实上,在这种情况下, P_2 所对应的物体往往已离开视频画面或在 k 帧未被检测到,而 O_2 是刚进入视频或在 k 帧刚被检测到的物体.因此,需要加入检验条件:

$$\begin{cases} T_i(P_i, O_j) \text{ 成立} & \text{Distance}(P_i, O_j) < T \\ T_i(P_i, O_j) \text{ 不成立} & \text{Distance}(P_i, O_j) \geq T \end{cases} \quad (11)$$

取 $T = \text{traj_length} \times \max \frac{(P_{iw}, P_{ih})}{2}$

式中: traj_length ——跟踪轨迹 T_i 累计无观测物体更新的帧的次数;

P_{iw}, P_{ih} ——已有轨迹斑块 P_i 的宽和高;

P_i ——轨迹 T_i 在当前帧的估计值;

O_j ——当前帧观测到的物体;

$T_i(P_i, O_j)$ —— P_i 和 O_j 在同一物体轨迹上.

这一修正,即式(11)的成立条件为:

(1) 假设视频中所有物体不会在瞬间突然提速或降速;

(2) 物体每秒行进距离小于每秒帧数与其自身长度的乘积的 50%,即假设视频为 15 fps, 4 m 长的车辆速度小于 30 m/s.

将数据关联结果代入式(5),可求得 k 帧最优轨迹估计作为跟踪结果显示.同时,需要对卡尔曼增益和 k 帧最优估值方差进行更新.

4 实验结果与分析

为了验证本方法实际场景应用的有效性,可针对多个不同场景视频序列进行跟踪实验.实验中,对每个跟踪物体均标明跟踪号,以此确定跟踪结果的正确性.

图 5 为车辆视频的跟踪结果.图 5a 和图 5c 中的黑色矩形框标识出跟踪到的物体,其右上角数字为跟踪物体号.图 5b 和图 5d 中白色斑块显

示背景检测算法提取到的前景物体.

实验环境: 双核 2.66 GHz CPU 3 GB 内存, 帧

图像大小为 352×288 像素, 每帧视频 MHT 算法处

理时间小于 1 ms, GMMs + MHT 处理时间为 47 ms.

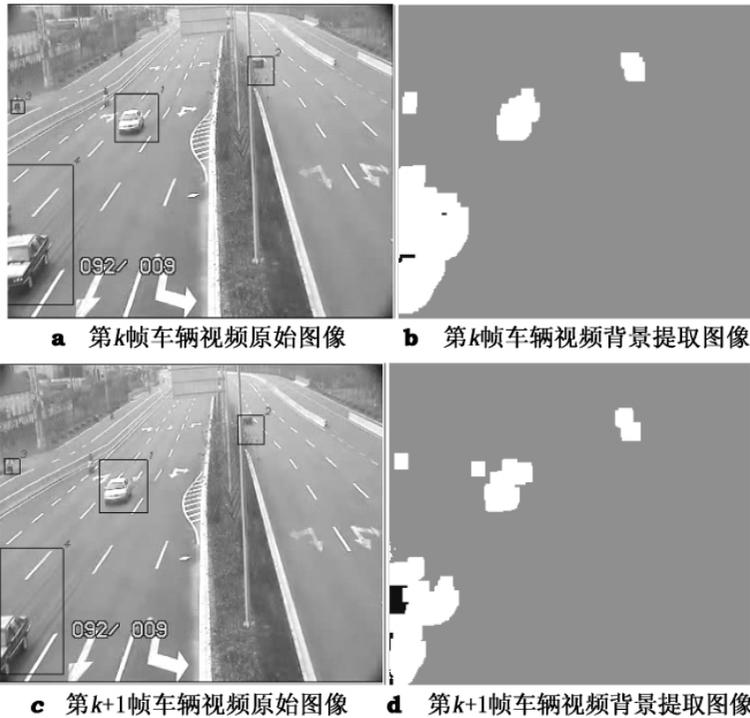


图5 连续两帧视频图像车辆跟踪效果

5 结束语

本文经过对多例实际场景视频的跟踪实验, 均获得良好的跟踪效果, 尤其对具有良好背景分割效果的视频得到了准确的跟踪结果. 实验表明, 该算法具有良好的鲁棒性. 但缺陷是对背景提取算法的过分依赖. 错误的跟踪主要来自于前景的错误提取, 这也是今后系统改进的重点之一.

参考文献:

[1] SAMUEL S B, RAYTHEON. Multiple hypothesis tracking for

multiple target tracking [J]. Aerospace and Electronic Systems Magazine, IEEE, 2004, 19(1): 5-18.

[2] ALPER Yilmaz, OMAR Javed, MUBARAK Shah. Object tracking: a survey [J]. ACM Computing Surveys, 2006, 38(4): 13.

[3] MASAMICHI Kojima, HIROSHI Kameda. A study of target tracking using track-oriented multiple hypothesis tracking [J]. SICE '98, 1998: 933-938.

[4] DAVID A Forsyth, JEAN Ponce. Computer vision: a modern approach [M]. Pearson Education, 2002: 534-549.

[5] REID D B. An algorithm for tracking multiple targets [J]. IEEE Transactions on Automatic Control, 1979, 24(6): 843-854.