

改进联合概率数据关联的视频多目标快速跟踪

万 琴^{1,2}, 王耀南², 袁小芳²

(1. 湖南工程学院 电气信息学院, 湖南 湘潭 411104; 2. 湖南大学 电气与信息工程学院, 湖南 长沙 410082)

摘要: 针对监控范围较大、目标外观特征少的视频多目标数据关联及跟踪问题, 本文仅利用目标运动特征, 提出了一种基于联合概率数据关联(joint probabilistic data association, JPDA)的复杂情况下视频多目标快速跟踪方法. 首先采用murty算法求JPDA的最优 K 个联合事件, 大大降低了计算复杂度; 然后根据JPDA的关联概率讨论目标的运动情况, 分析在多目标新出现、遮挡、消失、分离(前景检测存在目标碎片)等复杂情况下当前帧量测与跟踪目标的数据关联问题, 获取复杂运动的多目标跟踪轨迹. 在多个监控视频上的实验结果表明, 该方法能大大提高跟踪性能, 实现复杂情况下的视频多目标快速跟踪.

关键词: 视频监控; 多目标跟踪; 联合概率数据关联; 复杂运动

中图分类号: TP273 **文献标识码:** A

Tracking multiple video objects based on improved joint probabilistic data association

WAN Qin^{1,2}, WANG Yao-nan², YUAN Xiao-fang²

(1. College of Electrical & Information Engineering, Hunan Institute of Engineering, Xiangtan Hunan 411104, China;

2. College of Electrical and Information Engineering, Hunan University, Changsha Hunan 410082, China)

Abstract: For the data association of video objects having little distinguishable features in large-scale monitoring scenes, we present a method for tracking multiple video objects in real-time based on the joint probabilistic data association (JPDA), in which the motion features of the objects are incorporated. First, the k -best joint events are computed by the Murty algorithm to reduce the complexity, and then, the motion situations of objects are analyzed by the association probability of JPDA. When objects are entering and exiting the field of view, merging and splitting (objects are detected as fragmented parts), the data association algorithm acquires the tracking trajectories of the objects. Experiments demonstrate the feasibility and performances of the proposed approach.

Key words: visual surveillance; multiple objects tracking; joint probabilistic data association; complex motion

1 引言(Introduction)

在多运动目标视频监控中, 需要根据检测结果进行数据关联来匹配、识别出连续帧间的不同运动目标, 从而实现多目标的跟踪. 在监控范围较大、运动目标外观特征相似或目标区域较小的情况下, 数据关联仅能依靠目标的运动特征完成, 而传统的数据关联方法应用于视频多目标跟踪还存在很多问题. 本文将针对视频监控中的监控范围较大、目标外观特征少、运动复杂的情况, 讨论如何实现多目标数据关联, 识别、跟踪不同运动目标, 获取运动轨迹.

目前解决多运动目标的数据关联问题, 主要的方法有: 1) 文献[1~3]提出根据最近邻方法, 计算落入跟踪门限内的量测, 判定目标的运动情况, 并根据有效量测直接估计、更新当前运动状态, 这种方法

计算量小, 但在目标数目较多、运动情况复杂时, 抗干扰能力差, 容易产生错误关联; 2) 杂波环境下多目标数据关联方法如联合概率数据关联(joint probability data association, JPDA)、多假设跟踪(multiple hypothesis tracking, MHT), 目前多用于红外、雷达小目标或机动目标跟踪系统, 如文献[4, 5], 虽然和视频监控跟踪有相似性, 但须满足一对一关联的约束, 而在视频监控跟踪系统中, 多运动目标常发生新出现、消失、遮挡、分离等复杂运动情况, 即出现一对多或多对一的关联情况, 因此需进一步研究这些经典的多目标数据关联方法如何应用于视频多目标跟踪; 3) 采用优化算法分析当前帧检测区域与跟踪目标的最优关联, 如文献[6, 7]采用的图优化、赋权二分图等方法, 但此类方法通常需要获取遮挡或分离区域外观特征进行匹配优化计算, 在目标区域小、外

观信息少的情况下,难以得到目标外观特征,关联准确性低。

本文针对监控范围较大、目标外观特征少的视频多目标跟踪,仅采用目标的运动特征,基于JPDA算法实现复杂情况下的多目标跟踪,提出采用murty算法快速计算JPDA算法中最优 K 个联合事件,避免在目标数目多的情况下产生关联匹配“组合爆炸”;根据JPDA算法的关联概率分析多目标跟踪中目标新出现、消失、遮挡、分离(前景检测不准确造成目标碎片)等复杂情况下数据关联及跟踪目标状态估计,实现视频多目标实时跟踪,并获取跟踪轨迹。

2 目标检测及状态选取(Object detection and state selection)

采用文献[8]提出的背景检测方法,将检测出的目标用其外接矩形框表示,由于本文针对监控范围较大、目标外观特征少的情况下跟踪目标,因此仅能利用目标运动特征进行跟踪。本文选取目标外接矩形框中心点作为运动状态估计对象,由于在监控视频中,通常相邻帧间时间间隔 Δt 很小,目标运动可近似认为是匀速运动,该线性系统状态方程和测量方程为^[9]

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{w}_k, \quad (1)$$

$$\mathbf{z}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k. \quad (2)$$

式(1)中: \mathbf{w} 为系统噪声反映线性系统模型的正确程度,具有均值为零的高斯分布,其协方差矩阵为 \mathbf{Q} ; \mathbf{v} 为观测噪声,是均值为零的白噪声序列,和 \mathbf{w} 互不相关,其协方差矩阵为 \mathbf{R} 。同时,定义状态向量 $\mathbf{x}_k = (p_x, v_x, p_y, v_y)^T$ 和观测向量 $\mathbf{z}_k = (p_x, p_y)^T$,其中 p_x, v_x, p_y, v_y 分别表示第 k 帧目标外接矩形框中心点 x 方向坐标值、速度值及 y 方向坐标值、速度值,则根据匀速运动动力学方程得到系统状态转移矩阵 \mathbf{A}, \mathbf{C} 为

$$\mathbf{A} = \begin{bmatrix} 1 & \Delta t & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (3)$$

3 传统JPDA算法的基本步骤及存在的问题(Basic steps of traditional JPDA algorithm and existing problems)

数据关联是多目标跟踪的关键问题,即需确定传感器接收到的量测信息和目标源的对应关系,当目标较多且相互靠近,数据关联过程十分复杂^[10]。JPDA算法的基本思想是对所有可能的目标关联解进行搜索,并在此基础上计算出最佳关联概率,从而得到目标的状态估计。下面根据文献[10]阐述的JPDA算法基本原理,分析采用传统JPDA算法跟踪视频多目标的基本步骤及存在的问题。图1所示为

采用传统JPDA算法跟踪视频目标的基本流程,主要包括以下几个基本步骤:

Step 1 产生确认矩阵。

用于表示当前帧的有效量测和多目标跟踪门之间的关系,量测即为当前帧图像采用背景检测方法得到的检测信息,目标跟踪门则根据上一帧跟踪结果计算。设跟踪门限为 t_g ,如滤波残差向量的范数 d_k^2 小于门限,则判定当前帧对应的量测有效,

$$d_k^2 \leq t_g, \quad (4)$$

$$d_k^2 = \tilde{\mathbf{z}}_{k|k-1}^T \mathbf{S}_k^{-1} \tilde{\mathbf{z}}_{k|k-1}, \quad (5)$$

式中 $\tilde{\mathbf{z}}_{k|k-1}, \mathbf{S}_k$ 分别是Kalman滤波器得到的滤波残差向量、残差协方差矩阵。则确认矩阵 Ω 表示当前帧有效量测与目标跟踪门间的关系,定义为

$$\Omega = \{\omega_j^t\}, \quad (6)$$

其中: $j = 1, 2, \dots, m, t = 0, 1, 2, \dots, N$ 。确认矩阵 Ω 的行表示当前帧的有效量测,列表示跟踪目标, $t = 0$ 表示虚警,对应 Ω 第1列的所有元素为1。

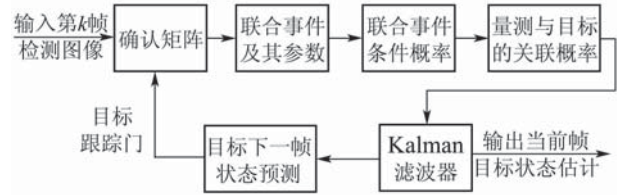


图1 基于传统JPDA算法的视频多目标跟踪基本流程
Fig. 1 Basic flow of tracking multiple video objects based on traditional JPDA algorithm

Step 2 产生所有联合(可行)事件及其参数。

为计算每一个量测与可能对应的各目标间的相互关联的概率,由确认矩阵 Ω 分析得到 k 时刻所有联合事件: $\theta_k = \{\theta_{k,i}\}_{i=1}^{n_k}$, n_k 表示集合 θ_k 中元素的个数,其中第 i 个联合事件为 $\theta_{k,i} = \prod_{j=1}^{m_k} \theta_{k,i}^{j,t_j}$,表示当前帧 m_k 个量测与各个目标匹配的一种可能。而表示第 j 个量测与目标 t 关联的事件为 $\theta_k^{j,t}$,称为关联事件。

再依据产生联合事件的两个基本假设^[10],对确认矩阵 Ω 拆分,得到与联合事件对应的可行矩阵

$$\hat{\Omega}(\theta_{k,i}) = [\hat{\omega}_j^t(\theta_{k,i})], \quad (7)$$

其中 $\hat{\omega}_j^t(\theta_{k,i})$ 描述在第 i 个联合事件中,量测 j 是否源于目标 t 。根据可行矩阵,得到联合事件 $\theta_{k,i}$ 中的参数:假量测数 $\Phi(\theta_{k,i})$ 、量测关联指示器 $\tau_j(\theta_{k,i})$ 及目标检测指示器 $\delta_t(\theta_{k,i})$ 。

Step 3 应用Bayes法则,计算 k 时刻所有量测的联合事件的条件概率,

$$P(\theta_{k,i} | \mathbf{Z}^k) = \frac{\Phi(\theta_{k,i})!}{cV^{\Phi(\theta_{k,i})}} \prod_{j=1}^{m_k} (A_{k,j})^{\tau_j(\theta_{k,i})} \times \prod_{t=1}^n (P_D^t)^{\delta_t(\theta_{k,i})} (1 - P_D^t)^{1 - \delta_t(\theta_{k,i})}, \quad (8)$$

其中: c 为归一化常数, V 表示跟踪门体积, $\Lambda_{k,j}$ 表示滤波残差似然函数, P_D^t 表示目标 t 的检测概率.

Step 4 计算量测与目标 t 的关联概率 $\beta_k^{j,t}, \beta_k^{0,t}$.

$$\begin{cases} \beta_k^{j,t} = P(\theta_k^{j,t} | Z^k) = \\ \sum_{i=1}^{n_k} P(\theta_{k,i} | Z^k) \hat{\omega}_j^t(\theta_{k,i}), j \in [1, m_k], \\ \beta_k^{0,t} = 1 - \sum_{j=1}^{m_k} \beta_k^{j,t}. \end{cases} \quad (9)$$

Step 5 通过Kalman滤波器, 得到目标状态估计及协方差阵.

目标 t 的状态估计

$$\hat{x}_{k|k}^t = \sum_{j=0}^{m_k} \beta_k^{j,t} \hat{x}_{k|k,j}^t. \quad (10)$$

目标 t 的状态估计协方差阵为

$$P_{k|k}^t = P_{k|k-1}^t - (1 - \beta_k^{0,t}) K_k^t S_k^t (K_k^t)^T + \sum_{j=0}^{m_k} \beta_k^{j,t} (\hat{x}_{k|k,j}^t (\hat{x}_{k|k,j}^t)^T - \hat{x}_{k|k}^t (\hat{x}_{k|k}^t)^T), \quad (11)$$

$$P_{k|k,0}^t = P_{k|k-1}^t, \quad (12)$$

其中: K_k^t, S_k^t 分别为目标 t 在 k 时刻的Kalman滤波增益矩阵、残差协方差矩阵.

在以上传统JPDA算法跟踪视频多目标的基本步骤中, 主要存在两个问题:

1) Step 2根据确认矩阵产生联合事件时, 如目标与量测较多, 对应的联合事件数目呈指数级增长, 计算复杂度非常高, 因此需进行合理简化, 降低计算复杂度, 才能应用于实际视频跟踪系统中;

2) 由于JPDA算法产生联合事件需要满足两个基本假设: 每一个量测有唯一的源; 对于一个给定目标, 最多有一个量测以其为源. 而在多目标视频跟踪中, 目标运动情况复杂, 如新目标出现、消失、遮挡、分离(前景检测造成的目标碎片), 量测、目标不满足“一对一”的关联约束, 传统JPDA算法跟踪误差变大, 最终会因累积误差过大而跟踪失败, 因此需对传统JPDA算法进行改进, 使之适用于复杂情况的视频多目标跟踪.

下面将分析在视频多目标跟踪中如何解决上述问题.

4 基于murty算法快速计算JPDA联合事件 (Fast computing JPDA joint events based on murty algorithm)

传统JPDA算法中对确认矩阵拆分得到联合事件对应的可行矩阵, 文献[11]给出可行矩阵数目的计算公式:

$$1 + \sum_{l=0}^{T-1} \frac{T!m!}{l!(T-l)!(m-T+l)!}, \quad (13)$$

其中: m 表示有效量测数, T 表示目标数, 式中第1项是对第一列全为1的可行矩阵 $\hat{\Omega}(\theta_{k,1})$ 的计数, 第2项

计算除可行矩阵 $\hat{\Omega}(\theta_{k,1})$ 外所有可行矩阵的数目. 从上式可见, 当 m, T 数目很大时, 可行矩阵数目及其对应的联合事件数目呈指数级增长, 产生“组合爆炸”现象, 如 $m = 10, T = 10$, 可行矩阵数达到234662231^[11], 计算复杂度很高, 因此需要分析如何快速计算可行矩阵及其对应的联合事件.

从式(10)可知, 对目标的状态估计主要取决于量测与目标的关联概率 $\beta_k^{j,t}$ 的大小, 而从计算 $\beta_k^{j,t}$ 的式(9)可见, 关联概率大小主要是由较大概率 $P(\theta_{k,i} | Z^k)$ 对应的联合事件决定的.

如图2所示, 当前帧两个量测 z_1, z_2 与两个目标的跟踪门形成确认矩阵 Ω :

$$\Omega = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad (14)$$

则联合事件对应的可行矩阵共有4个, 分别为

$$\begin{cases} \hat{\Omega}(\theta_{k,1}) = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \hat{\Omega}(\theta_{k,2}) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \\ \hat{\Omega}(\theta_{k,3}) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \hat{\Omega}(\theta_{k,4}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \end{cases} \quad (15)$$

在第6节实验分析的监控视频1中, 根据上述可行矩阵, 可得各个联合事件的条件概率为

$$\begin{cases} P(\theta_{k,1} | Z^k) = 0.0000, P(\theta_{k,2} | Z^k) = 0.0001, \\ P(\theta_{k,3} | Z^k) = 0.9999, P(\theta_{k,4} | Z^k) = 0.0000. \end{cases} \quad (16)$$

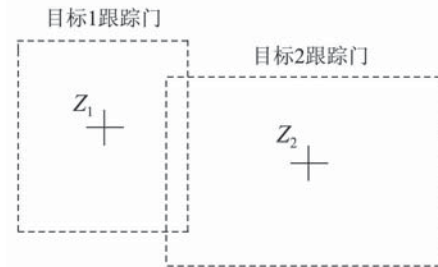


图 2 第 k 帧量测与目标跟踪门的关系
Fig. 2 Association of measurements and tracking gates of objects in the k -th frame

由式(9)得到量测与目标的关联概率矩阵为

$$\beta_k = \begin{bmatrix} 1.0000 & 0 \\ 0 & 0.9999 \\ 0 & 0.0001 \end{bmatrix}, \quad (17)$$

其中: 行表示跟踪目标, 列表示量测, 最后一行为虚量测. 从式(16)(17)可知, 图2所示的量测、目标间的关联概率 $\beta_k^{j,t}$ 的大小主要由条件概率较大的联合事件 $\theta_{k,2}, \theta_{k,3}$ 决定.

可见, 如果要避免产生所有由确认矩阵 Ω 得到的联合事件, 只须找出所有联合事件中对应条件概率最大的 K 个联合事件, 而那些对应较小后验概率的联合事件可忽略不计, 从而在保证关联概率计算正

确的情况下,大大降低计算复杂度.联合事件的数目 K 取值越大,计算结果越接近关联概率真实值,但是计算复杂度也会增加,实际应用中, K 值可根据系统计算能力和需达到的实时性高低共同确定,在本文实验中, K 取动态值,设定为当前帧跟踪目标数和量测数之和.

为了找出 K 个最大关联概率对应的联合事件,本文采用文献[12]提出的murty方法,该方法的基本思想是将给定的一个问题及解,划分为几个将解空间对应划分的子问题,通过找出子问题的解来获取原问题的解.因此,对应于获取确认矩阵 Ω 的最佳 K 个匹配关系,即认为确认矩阵 Ω 包含了所有可能的匹配解,每一次迭代找到一个最优匹配解,并在确认矩阵中删除这个最优匹配,下一次迭代再求一个最优匹配,如此循环 K 次,即得到确认矩阵 Ω 中最佳 K 个匹配关系.可见,问题的关键是求取每次迭代中的最优匹配.

最优匹配问题的求解是图论中的经典问题,通常采用赋权二分图^[13]描述待匹配的集合,如图2对应的确认矩阵可用如图3的赋权二分图描述.

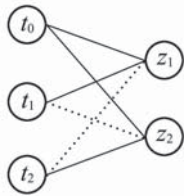


图3 图2对应的赋权二分图

Fig. 3 Weighted bipartite graph of fig.2

图3中左边节点表示跟踪目标,右边节点表示当前帧量测, t_0 表示虚警,连接目标和量测的线段表示一个可能的分配,实线对应权值为1,虚线对应权值为0.采用解决线性分配问题的匈牙利算法^[13]得到赋权二分图中 t, z 的最优匹配.

由于匈牙利算法要求赋权二分图的左、右节点数相同,因此如跟踪目标数(包括虚警)、量测数不一致时,采用增加虚目标或虚量测的方式,使左、右节点一致,如在图3中,量测少于跟踪目标数,则增加一个虚量测,虚量测与各目标间的权值为0,得到图4所示的赋权二分图,最后在每次迭代求得的最优解中删去增加的虚目标或虚量测. murty算法调用 K 次匈牙利算法即可找到对应于确认矩阵 Ω 的 K 个最优联合事件.

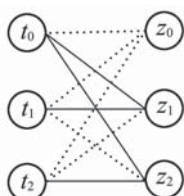


图4 图3增加一虚量测

Fig. 4 Adding a virtual measurement to fig.3

根据文献[12],由于匈牙利算法的时间复杂度为 $O(N^3)$ (N 表示赋权二分图中左、右节点中最大节点数),采用匈牙利算法实现的murty方法的最坏时间复杂度为 $O(KN^4)$,当量测和目标较多时,大大降低了计算可行矩阵及其联合事件的计算复杂度,避免了产生“组合爆炸”问题.

5 目标复杂运动分析及修正状态估计 (Analyze object complex motion and correct state estimation)

传统JPDA算法产生可行事件需满足量测、目标“一对一”的关联约束,而在视频监控中,多目标存在新出现、消失、遮挡、分离等运动情况,量测和目标的关联情况复杂,必须对传统JPDA算法进行改进,才能实现复杂运动情况下的视频多目标跟踪.

在第3节传统JPDA算法估计目标状态的式(10)中, $\hat{x}_{k|k,j}^t$ 表示由第 j 个量测得到的对目标 t 的状态估计由式(18)得到:

$$\hat{x}_{k|k,j}^t = \hat{x}_{k|k-1}^t + K_k(z_k - C_k \hat{x}_{k|k-1}^t), \quad (18)$$

其中: K_k 是Kalman增益矩阵, $\hat{x}_{k|k-1}^t$ 是一步预测状态向量:

$$\hat{x}_{k|k-1}^t = A_k \hat{x}_{k-1|k-1}^t. \quad (19)$$

式(19)中 $\hat{x}_{k-1|k-1}^t$ 表示目标 t 在上一帧($k-1$ 时刻)的状态估计值.

将式(18)(19)代入式(10)得

$$\hat{x}_{k|k}^t = \sum_{j=0}^{m_k} \beta_k^{j,t} \hat{x}_{k|k,j}^t = \sum_{j=0}^{m_k} \beta_k^{j,t} (A_k \hat{x}_{k-1|k-1}^t + K_k(z_k - C_k A_k \hat{x}_{k-1|k-1}^t)). \quad (20)$$

可见,当前帧状态估计 $\hat{x}_{k|k}^t$ 是由 $m_k + 1$ (包括虚量测)个量测得到的状态估计.根据式(20),要使传统JPDA得到正确的视频目标的状态估计 $\hat{x}_{k|k}^t$,须满足两个条件:

条件1 跟踪目标在上一帧的状态估计 $\hat{x}_{k-1|k-1}^t$ 必须存在;

条件2 量测与目标的关联概率 $\beta_k^{j,t}$ 准确可靠.

从以上分析可见,关联概率 $\beta_k^{j,t}$ 对目标的状态估计非常重要,由关联概率构成的关联矩阵 β_k (式(21))中前 m_k 行的关联概率 $\beta_k^{j,t}$ ($1 \leq j \leq m_k$)反映第 j 个有效量测与目标 t 的关联情况,而从式(9)中可知,关联矩阵 β_k 中的最后一行关联概率 $\beta_k^{0,t}$ 是由前 m_k 行的关联概率 $\beta_k^{j,t}$ 计算得到,表示没有任何量测源于目标 t 的概率.因此,本文选取关联矩阵 β_k 中前 m_k 行的关联概率构成新的关联矩阵 β'_k (式(21)), β'_k 的行表示当前帧有效量测,列表示上一帧的跟踪目标.通过关联矩阵 β'_k 分析目标的5种运动情况:目标正常、新出现、消失、遮挡、分离,判定在各运动情况下能否通

过式(20)得到准确有效的状态估计, 如存在较大误差, 则需修正状态估计.

$$\beta_k = \begin{bmatrix} \beta_k^{1,t_1} & \cdots \\ \vdots & \vdots \\ \beta_k^{m_k,t_1} & \cdots \\ \beta_k^{0,t_1} & \cdots \end{bmatrix}, \beta'_k = \begin{bmatrix} \beta_k^{1,t_1} & \cdots \\ \vdots & \vdots \\ \beta_k^{m_k,t_1} & \cdots \end{bmatrix}. \quad (21)$$

1) 正常. 如图2所示, 当目标 t 跟踪门内有且只有一个有效量测 z_j , 关联概率矩阵量测 z_j 所在行有唯一大于0的元素 t , 并且该元素所在列也只有该元素 t 唯一大于0时, 则当前量测 z_j 与元素 t 所在列对应的上一帧跟踪目标 t 关联, 认为当前帧目标 t 运动“正常”, 目标帧间状态稳定, 近似为线性运动, 因此可采用式(20)计算目标 t 的状态估计, 无须修正.

2) 新出现. 由于是当前帧出现的新目标, 没有上一帧的状态估计值 $\hat{x}_{k-1|k-1}^t$, 不满足状态估计的条件1, 无法通过式(20)计算当前帧的状态估计值. 因此需判定那些量测是新目标, 并确定初始跟踪状态, 判定新目标包括两个步骤:

① 初步判定. a) 如图5(a)所示, 当前帧量测 z_j 在所有目标的跟踪门外, 认为 z_j 可能代表新目标; b) 如关联概率矩阵中量测 z_j 所在行的元素全为0, 则当前帧量测 z_j 与所有上一帧目标无关联, 认为 z_j 可能代表新目标;

② 确认为“新目标”. 由于新目标可能是目标消失一段时间重新出现, 则该目标不能认为是新目标, 而应恢复对该消失目标的重新跟踪. 因此, 将初步判定为“新出现”的目标与前 k_s 帧(实验中设 $k_s = 5$)内所有状态为“消失”的目标进行第3节中Step 1~Step 5关联, 如能满足“正常”运动状态的条件, 则该新目标实际为“消失”目标重新出现, 则更新此目标状态为“正常”; 如该目标不能与任一“消失”目标关联, 则确认此目标为“新出现”.

通过以上分析, 判定出哪些量测是新目标, 新增加为跟踪对象, 并用其当前量测 $z_j(z_j = [p_x \ p_y]^T)$ 作为新出现目标的初始状态中的位置向量, 速度向量设为0, 即得到新目标的初始状态估计 $\hat{x}_{k|k}^t = [p_x \ 0 \ p_y \ 0]^T$.

3) 消失. 当目标在当前帧消失时, 即使跟踪门内无有效量测时, 但根据式(9)可得 $\beta_k^{0,t} = 1$, 同时目标存在上一帧的状态估计值 $\hat{x}_{k-1|k-1}^t$, 因此仍可根据式(20)得到消失目标在当前帧的状态估计, 显然与目标已消失的实际运动情况不符. 因此须根据量测与目标的关联情况, 判定目标是否消失, 以终止消失目标的跟踪.

如关联概率矩阵 β'_k 中上帧跟踪目标 t 所在列的元

素全为0, 则表明当前帧无量测与上一帧目标 t 关联, 目标 t 在当前帧可能消失, 如连续3帧判定为消失, 则确认该目标已消失, 终止该目标的跟踪.

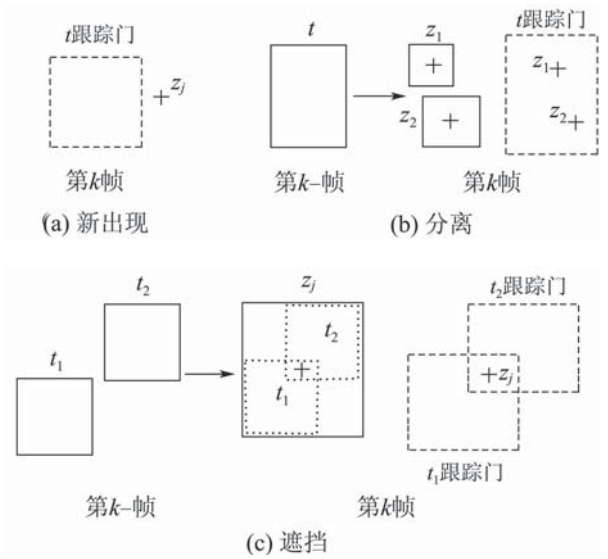


图 5 不同运动情况下量测与目标跟踪门的关系

Fig. 5 Association of measurements and tracking gates of objects under different motion situations

4) 遮挡. 多个目标在当前帧相互遮挡, 只能检测为一个前景区域, 如图5(c)所示, 第 $k - 1$ 帧的两个运动目标 t_1, t_2 在当前帧相互遮挡, 检测为一个前景区域 z_j , 无法分别得到各个目标区域的量测值, 即仅得到遮挡区域中心点的量测值(在图5(c)中用“+”号表示), 该量测值在遮挡目标跟踪门内(图5(c)中目标跟踪门与量测的关系). 由于遮挡区域的中心点量测值与各遮挡目标实际中心点位置存在差异, 仅采用该量测与各跟踪目标的关联概率进行状态估计必然存在较大误差, 即不满足状态估计的条件2. 如当量测与某一目标关联概率较大, 而与其他目标关联概率非常小时, 传统JPDA算法只能得到关联概率较大的目标状态估计, 其他遮挡目标状态估计近似为0. 如表1所示, 遮挡目标 t_2 与量测 z_j 关联概率远远大于遮挡目标 t_1 与量测 z_j 的关联概率, 此时目标 t_1 状态估计近似为0(即其位置、速度近似为0), 与目标 t_1 的真实位置、速度误差很大, 因此需判定目标是否遮挡, 并重新修正每个遮挡目标的状态估计.

如关联概率矩阵 β'_k 中量测 z_j 所在行有多个大于0的元素(如表1所示), 则量测 z_j 与这些元素所在列表的上一帧跟踪目标关联, 即多个上一帧目标在当前帧发生遮挡, 并且该行大于0的元素个数就是遮挡目标的个数. 由于量测 z_j 表示遮挡区域在当前帧的位置, 遮挡目标 t_1, t_2 状态估计中的位置向量修正为量测 z_j 的值、速度向量修正为0, 即

$$\hat{x}_{k|k}^{t_1} = [p_x \ 0 \ p_y \ 0]^T, \hat{x}_{k|k}^{t_2} = [p_x \ 0 \ p_y \ 0]^T.$$

表1 遮挡情况下量测 z_j 所在行

Table 1 The row of measurement z_j under occlusion

t_1	t_2
2.7939e-8	0.6231

5) 分离. 由于目标检测结果无法做到完全准确, 因此可能造成检测到的目标不完整(即目标碎片), 认为目标在当前帧发生分离(见图5(b)所示), 第 $k-1$ 帧目标 t 在当前帧分离为两个检测区域 z_1, z_2 . 由于传统JPDA跟踪方法的本质正是给出跟踪目标与当前多个有效量测间的关联情况, 并据此得到目标状态估计, 而目标分离产生的多个目标碎片恰好提供了该跟踪目标的多个量测中目标跟踪门与量测的关系(见图5(b)所示), 满足状态估计的条件1, 2, 因此根据式(2)能准确有效地估计跟踪目标在当前帧的运动状态, 无须修正状态估计值.

如关联概率矩阵 β'_k 中上帧跟踪目标 t 所在列有多个大于0的元素, 则该列表示的上一帧跟踪目标 t 与这些元素所在行对应的多个量测关联, 表示上一帧跟踪目标 t 在当前帧分离, 如当前帧量测 z_1, z_2 在目标 t 的跟踪门内, 并且关联矩阵中目标 t 所在的列如表2所示.

表2 分离目标 t 所在列

Table 2 The column of separated object t

z_1	z_2
0.0005	0.0005

从表2中可见, 跟踪目标 t 在当前帧分离为量测 z_1, z_2 , 并且关联概率给出了量测 z_1, z_2 与目标 t 的关联性, 据此通过式(20)得到目标 t 的状态估计.

图6所示为本文提出的基于JPDA算法的视频多目标快速跟踪的流程图, 以下为对应于图6的程序伪

代码.

本文算法伪代码: 手动设定初始目标跟踪区域.

for 每个当前视频帧(时刻 k),

读入上一帧运动情况为未“消失”的所有跟踪

目标 T 、状态估计 $\hat{x}_{k-1|k-1}$,

读入当前帧检测到的目标量测 z_k ;

for $i = 1 : m$,

调用基本JPDA算法函数, 得到跟踪目标在当前帧 $\hat{x}_{k|k}^t$ 、关联概率矩阵 β'_k ,

调用判断目标运动情况函数, 获得目标运动情况的标记state;

if state=“新出现”,

读入前 k_s 帧内所有“消失”目标的状态估计,

调用基本JPDA算法, 得到“消失目标”的状态估计 $\hat{x}_{k|k}$ 与“新出现”目标的关联概率矩阵

β'_k ,

if 关联矩阵 $\hat{\beta}_k \neq 0$,

if “新”目标所在列有唯一元素为0,

“新”目标与“消失”目标关联, 状态估计为 $\hat{x}_{k|k}$,

else 关联概率矩阵 $\hat{\beta}_k = 0$,

“新”目标与所有“消失”目标不关联,

确认是新目标, 根据量测 z_k 修正初始状态,

if state=“消失”,

终止目标跟踪;

if state=“遮挡”,

根据量测 z_k 修正目标状态估计;

if state=“分离”, 目标状态估计为 $\hat{x}_{k|k}$;

if state=“正常”,

目标状态估计为 $\hat{x}_{k|k}$.

记录当前帧所有跟踪目标.

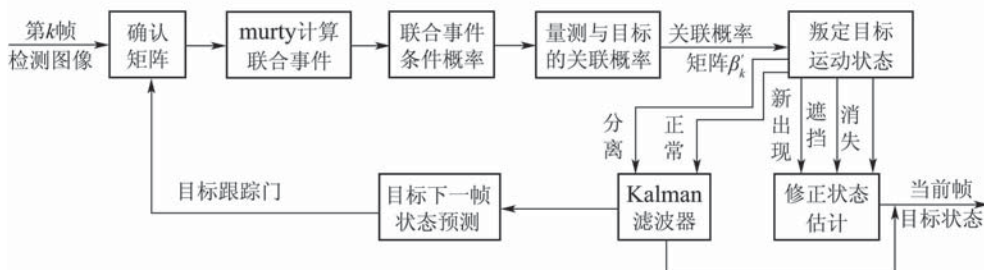


图6 基于JPDA算法的视频多目标跟踪流程图

Fig. 6 Flowchart of tracking multiple video objects based on JPDA algorithm

6 实验分析(Experiment analysis)

本文提出的方法在配置为奔4处理器(主频3.00 GHz)、1GB内存的台式计算机、MATLAB 7.1的编

程环境下对多个监控视频进行实验分析. 实验1的监控视频使用的是PETS-ECCV 2004数据库的视频, 选取第270帧到第503帧共233帧进行实验分

析, 图像分辨率为 384×288 , 监控场景中有两人不规则运动, 并且人体区域较小; 实验2是VS-PETS 2003标准数据库提供的一段205帧的足球比赛视频序列, 图像分辨率为 720×576 , 监控场景中有多个外观颜色相似的运动员, 目标运动复杂, 并且目标距离摄像机远、区域小。

图7所示为本文方法在监控视频1上的跟踪结果, 其中图7(e)~7(h)是采用前景检测算法^[8]得到

的检测图像, 矩形框表示检测为前景目标的区域, 数字表示跟踪目标的标号, “+”号表示当前帧目标状态估计的位置。从图7(e)(f)中可见在前景区域检测不完整、有目标碎片的情况下, 能正确跟踪目标。在图7(g)中目标出现相互遮挡、并有目标碎片时, 本文方法仍能有效跟踪不同目标, 并能在图7(h)所示的目标恢复为运动正常后, 准确识别、跟踪目标1, 2。

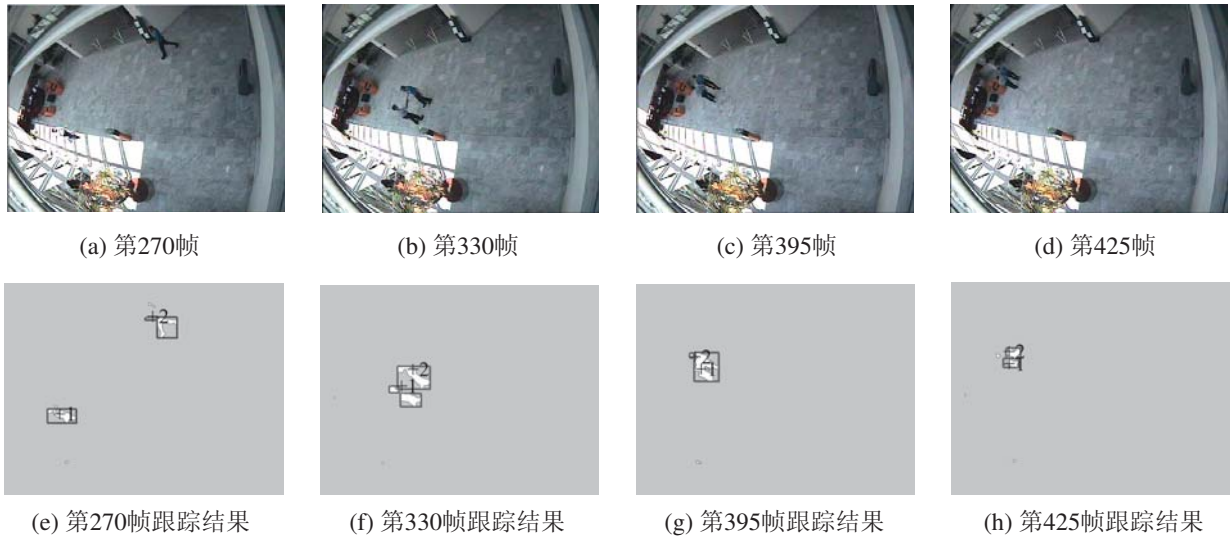


图 7 监控视频1的部分跟踪结果

Fig. 7 Partial tracking results in surveillant video 1

图8显示了目标在视频中的轨迹, 可见本文提出的算法能连续有效地跟踪多个目标。



图 8 监控视频1目标轨迹

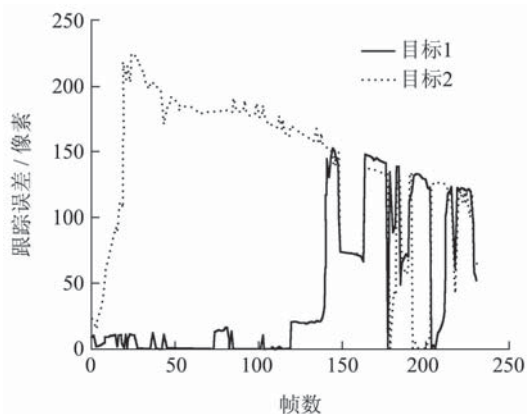
Fig. 8 Trajectories of objects in surveillant video 1

图9给出了不同方法在监控视频1中的跟踪误差, 跟踪误差是由目标中心真实坐标值与目标状态估计的位置值的欧式距离得到, 其中图9(a)~9(c)分别是采用传统JPDA算法^[10]、经典的实时视频多目标跟踪方法^[3]、本文算法跟踪目标1, 2的误差结果。监控视频1虽然只有两个目标, 但是外观有一定相似性, 场景存在光线等的变化, 因此检测结果存在目标碎片, 而且两目标运动出现遮挡等复杂情况, 由于文献^[10]的传统JPDA算法需要满

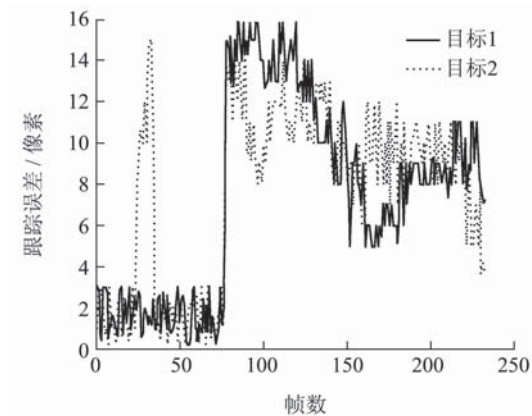
足“一对一”的关联约束, 文献^[3]中经典的实时跟踪算法采用简单的目标外观颜色、运动特征进行匹配, 均无法处理新目标出现、消失、遮挡、分离(前景检测造成的目标碎片)等复杂运动情况, 因此跟踪误差很大, 而本文提出的算法, 采用运动特征分析了不同情况下的跟踪策略, 在监控范围较大、目标外观相似的情况下, 能大大提高跟踪性能。由于文献^[3, 10]中的方法存在较大跟踪误差并且在累积, 而本文方法在每帧中能根据不同目标运动情况进行状态估计, 误差较小, 在跟踪100帧后, 由于目标越来越远离场景, 即目标区域越来越小, 本文提出的方法能实现目标的有效跟踪, 计算的目标状态中位置值与目标真实位置的差值必然越来越小, 接近于0。在图9中可见, 整个视频中本文方法的跟踪误差远远小于传统JPDA算法及文献^[3]的跟踪方法, 最大误差值仅不到4个像素, 同时表3给出了跟踪误差的统计量, 可见本文算法大大提高了跟踪精度, 并且每秒处理帧数(fps)达到了13fps, 能实现实时快速跟踪。

图10所示为本文方法在监控视频2上的跟踪结果, 矩形框、数字及“+”号含义与图7相同, 从

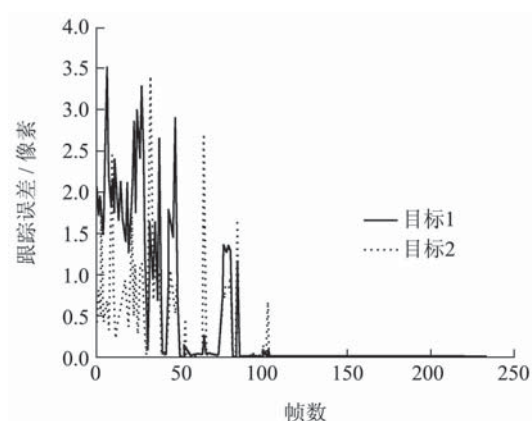
图10的跟踪结果可见, 本文提出的方法在目标检测不完整、目标数目多、外观相似且运动复杂的情况下, 能较好地实现目标跟踪并获取运动轨迹, 每个目标的运动轨迹见图11. 文献[10]的传统JPDA算法在监控视频2的实验中只能完成前4帧跟踪, 由于误差累积很大, 第5帧开始跟踪失败; 而文献[3]中经典的实时跟踪算法由于仅简单将目标跟踪门内的有效量测与该目标匹配, 并且需采用目标的外观颜色作为特征进行匹配, 无法实现对监控视频2中大量外观颜色相似、运动复杂的目标进行跟踪, 从第3帧开始就跟踪失败, 因此, 图12只显示了由本文方法得到的运动情况复杂的目标8、目标18的跟踪误差, 表4为跟踪误差的统计量及跟踪时间, 可见在目标数目多、运动情况复杂的情况下, 本文方法仍能有效跟踪多运动目标, 并且跟踪精度满足视频监控需求、跟踪实时快速.



(a) 传统JPDA算法^[10]



(b) 实时视频跟踪算法^[3]



(c) 本文算法

图9 不同方法在监控视频1中的跟踪误差
Fig. 9 Tracking errors of different algorithms for surveillant video 1

表3 监控视频1跟踪误差的统计量及跟踪速度

Table 3 Statistics of tracking errors and tracking time for surveillant video 1

算法	目标1		目标2		跟踪速度/fps
	均值	方差	均值	方差	
传统JPDA算法 ^[10]	43.1800	2.9386e+003	137.0330	3.4400e+003	0.7
实时跟踪算法 ^[3]	7.4824	23.7066	7.5867	18.6811	10
本文算法	0.4061	0.6534	0.2060	0.2134	13



(a) 第26帧



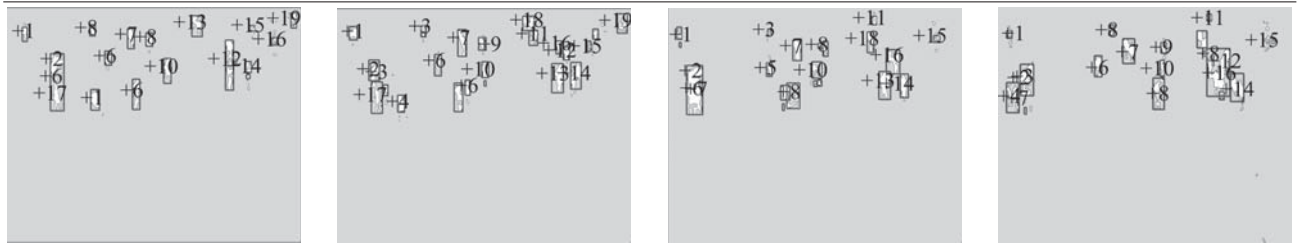
(b) 第61帧



(c) 第110帧



(d) 第150帧



(e) 第26帧跟踪结果 (f) 第61帧跟踪结果 (g) 第110帧跟踪结果 (h) 第150帧跟踪结果

图 10 监控视频2的部分跟踪结果

Fig. 10 Partial tracking results in surveillant video 2

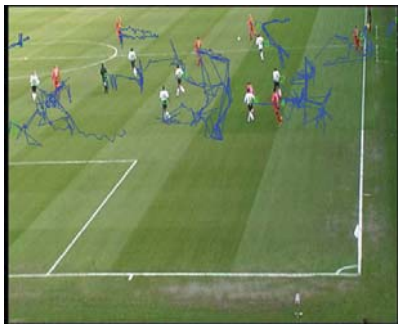
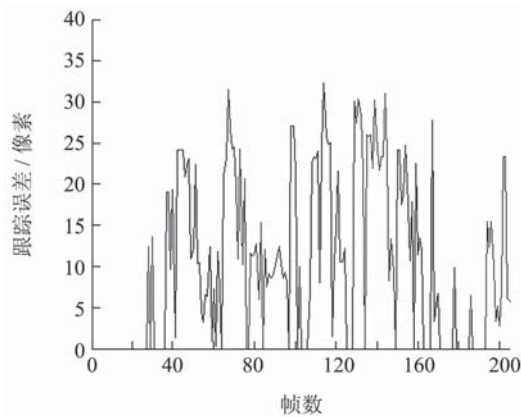
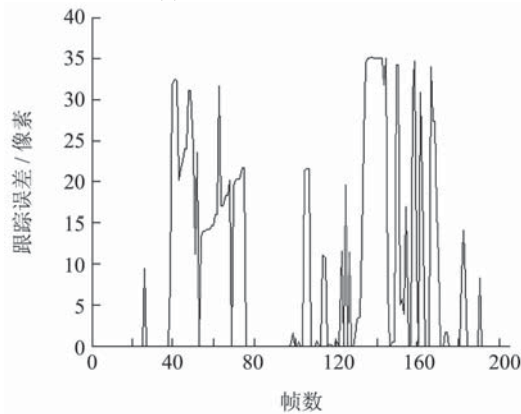


图 11 监控视频2目标轨迹

Fig. 11 Trajectories of objects in surveillant video 2



(a) 目标8的跟踪误差



(b) 目标18的跟踪误差

图 12 本文方法在监控视频2中的跟踪误差

Fig. 12 Tracking errors of the algorithm in this paper for surveillant video 2

表 4 监控视频2跟踪误差的统计量及跟踪速度
Table 4 Statistics of tracking errors and tracking time for surveillant video 2

算法	目标8		目标18		跟踪速度/fps
	均值	方差	均值	方差	
本文算法	10.0874	102.7518	8.239	141.6842	4

7 结论(Conclusions)

数据关联是多目标跟踪的关键问题, 本文针对监控范围较大、运动目标外观特征少的多目标数据关联及跟踪问题, 仅采用目标运动特征, 提出一种基于JPDA的视频多目标快速跟踪方法. 本文分析了传统JPDA算法应用于视频监控中跟踪多目标的基本步骤, 但由于目标、量测在目标新出现、消失、遮挡、分离(前景检测不准确造成的目标碎片)等复杂情况下无法一一对应, 同时在目标较多时, 关联匹配可能出现“组合爆炸”, 因此传统的JPDA算法不能直接应用于视频多目标数据关联及跟踪. 为避免在目标较多时的关联匹配呈指数增加, 本文采用murty算法求JPDA的最优 K 个联合事件, 能大大降低计算复杂度; 根据JPDA的关联概率分析目标的运动情况, 分别讨论了在目标正常、新出现、消失、遮挡、分离等复杂情况下当前帧量测与跟踪目标的数据关联、状态估计问题, 从而有效实现了视频多目标复杂运动的快速跟踪. 在标准数据库上的实验证明, 本文方法能在复杂场景中有效跟踪多个外观相似的视频目标, 跟踪精度比文献[10]的传统JPDA方法及文献[3]的经典实时视频跟踪方法大大提高, 并能实现实时快速跟踪.

由于本文方法是跟踪前景检测结果进行分析, 当复杂场景中前景检测精度太低、目标碎片很多的情况下, 会影响本文方法的有效性. 因此, 下一步工作希望能直接利用视频原图像分析目标运动特征, 并合理采用有效的外观特征, 进一步提高跟踪性能.

参考文献(References):

- [1] FUENTES L M, VELASTIN S A. People tracking in surveillance applications[J]. *Image and Vision Computing*, 2006, 24(11): 1165 – 1171.
- [2] CAVALLARO A, STEIGER O, EBRAHIMI T. Tracking video objects in cluttered background[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2005, 15(4): 575 – 584.
- [3] KAEWTRAKULPONG P, BOWDEN R. A real time adaptive visual surveillance system for tracking low-resolution colour targets in dynamically changing scenes[J]. *Image and Vision Computing*, 2003, 21(10), 913 – 929.
- [4] 李正周, 金钢, 董能力. 基于改进概率数据关联滤波的红外小目标运动目标跟踪[J]. 电子与信息学报, 2008, 30(4): 953 – 956. (LI Zhengzhou, JIN Gang, DONG Nengli. A novel method for tracking and recognizing infrared dim and small moving target based on modified probabilistic data associating filter[J]. *Journal of Electronics and Information Technology*, 2008, 30(4): 953 – 956.)
- [5] 耿峰, 祝小平. 一种改进的多传感器多目标跟踪联合概率数据关联算法研究[J]. 系统仿真学报, 2007, 19(20): 4671 – 4675. (GENG Feng, ZHU Xiaoping. Research of improved joint probabilistic data association algorithm for multisensor-multitarget tracking[J]. *Journal of System Simulation*, 2007, 19(20): 4671 – 4675.)
- [6] CONTE D, FOGGIA P, JOLION J M, et al. A graph-based, multi-resolution algorithm for tracking objects in presence of occlusions[J]. *Pattern Recognition*, 2006, 39(4): 562 – 572.
- [7] PARK S, AGGARWAL J K. Simultaneous tracking of multiple body parts of interacting persons[J]. *Computer Vision and Image Understanding*, 2006, 102(1): 1 – 21.
- [8] ELGAMMAL A, HARWOOD D, DAVIS L S. Non-parametric model for background subtraction[C] // *European Conference on Computer Vision*. Ireland: Springer, 2000: 751 – 767.
- [9] FORSYTH D A, Ponce J. 计算机视觉——一种现代方法[M]. 北京: 电子工业出版社, 2004: 323 – 324. (FORSYTH D A, PONCE J. *Computer Vision: A Modern Approach*[M]. Beijing: Electronic Industry Press, 2004: 323 – 324.)
- [10] 韩崇昭, 朱洪艳, 段战胜, 等. 多源信息融合[M]. 北京: 清华大学出版社, 2006: 297 – 302. (HAN Chongzhao, ZHU Hongyan, DUAN Zhansheng, et al. *Multi-Source Information Fusion*[M]. Beijing: Tsinghua University Press, 2006: 297 – 302.)
- [11] FISHER J L, CASASSENT D P. Fast JPDA multitarget tracking algorithm[J]. *Applied Optics*, 1989, 28(2): 371 – 376.
- [12] COX I J, HINGORANI S L. An efficient implementation of reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(2): 138 – 150.
- [13] 卢开澄, 卢华明. 图论及其应用[M]. 第2版. 北京: 清华大学出版社, 2005: 189 – 202. (LU K C, LU H M. *Graph Theory with Applications*[M]. 2nd edition. Beijing: Tsinghua University Press, 2005: 189 – 202.)

作者简介:

万琴 (1980—), 女, 讲师, 博士, 目前研究方向为模式识别、图像处理、视频目标检测与跟踪, E-mail: wanqin_10@126.com;

王耀南 (1957—), 男, 教授, 博士, 目前研究方向为模式识别与图像处理、智能控制理论与应用、智能机器人系统;

袁小芳 (1979—), 男, 讲师, 博士, 目前研究方向为智能控制理论与应用、优化计算、神经网络.