

一种改进的离散连续能量最小化多目标跟踪

张丽娟, 周治平

江南大学物联网技术应用教育部工程研究中心, 江苏 无锡 214122

摘要 针对离散-连续能量最小化(DCEM)方法在复杂场景中对轨迹分段或身份标签互换无法有效处理的问题,提出一种改进的 DCEM 多目标跟踪方法。该方法通过提取被跟踪目标的多特征融合外观向量,利用不同目标间外观特征向量的欧氏距离设计轨迹的外观约束项,处理身份标签互换问题;通过计算相邻时空域内不同轨迹间的运动相似性和外观相似性,设计后处理过程,合并可能为同一轨迹的短轨迹,处理轨迹分段问题。实验结果表明,平均跟踪准确度提高 3.6%,平均跟踪精度提高 2.5%,并且身份标签互换和轨迹分段情况得到大幅改善,该方法具有更精确更稳定的跟踪能力。

关键词 机器视觉; 多目标跟踪; 能量最小化; 身份标签互换; 轨迹分段

中图分类号 TP391 **文献标识码** A

doi: 10.3788/LOP54.111502

An Improved Discrete-Continuous Energy Minimization for Multi-Target Tracking

Zhang Lijuan, Zhou Zhiping

Engineering Research Center of Internet of Things Technology Applications, Ministry of Education, Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract Aiming at the problems that the discrete-continuous energy minimization (DCEM) method cannot effectively deal with the trajectory segmentation and identity switch in complex scene, an improved DCEM multi-target tracking method is put forward. In order to solve identity switch, this method extracts the multi-feature fusion appearance vector of the tracked target, and uses the Euclidean distance between appearance vectors from different targets to design the constraint function for trajectories. Then this method designs the post-processing process for trajectory segmentation problem by merging the short tracklets which have high similarity of motion and appearance in adjacent spatial-temporal neighborhood. Experimental results indicate that the average tracking accuracy and average tracking precision can increase by 3.6% and 2.5%, respectively. Furthermore, the problems of trajectory segmentation and identity switch can be greatly improved. The proposed method has better robustness and the tracking accuracy.

Key words machine vision; multi-target tracking; energy minimization; identity switch; trajectory segmentation

OCIS codes 150.0155; 200.3050; 100.2000; 100.4999

1 引言

多目标跟踪是跟踪图像序列中的多个移动目标并随时间分配各个目标身份的问题,目标图像不完整、遮挡和复杂的目标运动等都会增加跟踪的难度。在发展初期,多目标跟踪主要是独立检测的跟踪,即在第一帧标定被跟踪目标的位置然后用滤波方法逐帧跟踪目标。但这种方法只能跟踪固定数量的目标,故其实用价值有限。随着运动目标检测技术的完善与发展^[1-3],基于检测的跟踪方法逐渐流行起来。

收稿日期: 2017-04-05; **收到修改稿日期:** 2017-06-04

基金项目: 中央高校基本科研业务费专项资金(JUSRP51510)

作者简介: 张丽娟(1993—),女,硕士研究生,主要从事音视频信号分析与处理方面的研究。

E-mail: 6151904003@vip.jiangnan.edu.cn

导师简介: 周治平(1962—),男,博士,教授,主要从事检测技术和自动化装置方面的研究。E-mail: zzp@jiangnan.edu.cn

基于检测的跟踪方法大致分递归和非递归两种。递归跟踪方法通过当前帧和过去帧的观测值来推断目标之后的状态^[4-7],但由于无法在跟踪过程中矫正目标位置,因此造成误差累积导致跟踪漂移。非递归跟踪方法的特点是所有的目标轨迹都在一定时间窗内通过联合估计获得。最近越来越多的研究者将最小费用网络流优化技术应用于多目标跟踪中,这种方法可以同时分析视频中多个对象的轨迹,并且可以对轨迹之间的相互关系进行建模。Zhang 等^[8]提出网络流框架,将最大后验数据关联问题映射到具有非重叠轨迹约束的费用流网络中,目标函数的全局最优解可以在最小成本流算法中找到。但由于其仅考虑不同轨迹之间位置关系约束,易造成遮挡目标的跟踪失败或轨迹分段等问题。Pirsiavash 等^[9]在上述框架下,提出了一种更快的贪婪算法,在成本网络中寻找最短路径,得到最优轨迹,运算速度大大提升。但其同样未对轨迹添加约束,遮挡等问题依旧存在。Butt 等^[10]采用轨迹图模型,考虑了目标的高阶属性,如目标的运动等,有效提高了跟踪精度与准确度。但对于位置相似、运动相似的不同目标无法做出有效区分,轨迹身份标签互换现象严重。Milan 等^[11]提出一种离散-连续能量最小化(DCEM)的方法,将数据关联的离散问题与轨迹估计的连续问题整合到一个能量函数中,并用目标运动模型对轨迹进行约束,采用带有身份标签的能量最小化算法求得最优解,得到具有较高跟踪精度与准确度的跟踪结果。但轨迹分段或身份标签互换问题仍不能得到有效解决。为此,本文提出一种改进的离散-连续能量最小化方法,考虑目标外观与运动对轨迹的影响,根据目标外观的相似性设计外观约束项,根据目标外观与运动的相似性设计后处理过程,避免轨迹分段与身份标签互换,实现准确的多目标跟踪。

2 相关研究

为了将连续轨迹估计和离散数据关联问题相统一,文献[11]采用连续曲线表示轨迹,用离散多标签条件随机场(CRF)进行建模。设计离散-连续能量函数后,用带有标签的能量最小化优化方法交替计算离散数据关联和连续轨迹估计,可以有效最小化 CRF 能量,得到最小能量、轨迹及检测对应的身份标签。离散-连续能量函数为

$$E(\Gamma, f) = \sum_{d \in D} E_{\text{det}}(f_d, \Gamma_d) + \sum_{(d, d') \in E_s} E_{\text{sm}}(f_d, f_{d'}) + \sum_{(d, d') \in E_\chi} E_{\text{det}}^\chi(f_d, f_{d'}) + \sum_{\Gamma_i \in \Gamma_f^*} E_{\text{traj}}(\Gamma_i) + \sum_{\Gamma_i, \Gamma_j \in \Gamma_f^*, i \neq j} E_{\text{traj}}^\chi(\Gamma_i, \Gamma_j), \quad (1)$$

式中 D 表示所有检测结果的集合, f 为对应的身份标签集合, $f_d \in f$ 表示为每个检测结果 $d \in D$ 分配的身份标签, Γ 为所有目标轨迹假设, $\Gamma_d \in \Gamma$ 为其中一条轨迹,由三次 B-样条曲线拟合表示, $\Gamma_f^* = \{\Gamma_i | \exists d \in D: f_d = i\}$ 为分配有效身份标签的活跃轨迹, E_{det} 项表示分配给轨迹的检测结果距离轨迹的距离, E_{traj} 项表示轨迹的一些如运动状态等性质, E_{traj}^χ 项表示轨迹间相互不重叠, E_s 与 E_χ 分别表示在 CRF 中随时间递进检测相连形成的边集和同一时间同一帧中不同检测相连形成的边集 E_{sm} 和 E_{det}^χ 分别表示时间平滑项和检测级互斥项。

离散-连续能量函数中的能量主要分为三部分:连续轨迹估计、离散数据关联和轨迹标签成本。连续轨迹估计是指在已知轨迹数据关联 f 时,轨迹估计问题可以看作求解(1)式中连续部分能量的最小值,即

$$E_f^{\text{tc}}(\Gamma) = \sum_{d \in D} E_{\text{det}}(f_d, \Gamma) + \sum_{\Gamma_i \in \Gamma_f^*} E_{\text{traj}}(\Gamma_i) + \sum_{\Gamma_i, \Gamma_j \in \Gamma_f^*, i \neq j} E_{\text{traj}}^\chi(\Gamma_i, \Gamma_j), \quad (2)$$

在给定身份标签 f 时有 $\arg \min_{\Gamma} E_f^{\text{tc}}(\Gamma) = \arg \min_{\Gamma} E(\Gamma, f)$ 。

数据关联通常是多目标跟踪中最具挑战性的方面,文献[11]中将其定义为多标签问题以便利用离散优化算法求解。数据关联问题在此可以定义为估计一个身份标签 f ,使其可以将每个检测 $d \in D$ 分配给 N 条轨迹假设 $\Gamma = \{\Gamma_1, \dots, \Gamma_N\}$ 中的一个,或者标记为异常值 φ ,表示为

$$E_f^{\text{da}}(f) = \sum_{d \in D} E_{\text{det}}(f_d, \Gamma_d) + \sum_{(d, d') \in E_s} E_{\text{sm}}(f_d, f_{d'}) + \sum_{(d, d') \in E_\chi} E_{\text{det}}^\chi(f_d, f_{d'}) + \sum_{\Gamma_i \in \Gamma_f^*} E_{\text{traj}}(\Gamma_i) + \sum_{\Gamma_i, \Gamma_j \in \Gamma_f^*, i \neq j} E_{\text{traj}}^\chi(\Gamma_i, \Gamma_j), \quad (3)$$

(3)式除了已知轨迹假设 Γ 外,与(1)式保持一致。

(1)式中 $E_{\text{traj}}(\Gamma_i)$ 为轨迹的标签成本项,包括角速度约束项、线速度约束项、轨迹持续性项和边界条件,定义如下:

$$E_{\text{traj}}(\Gamma_i) = E_{\text{ang}}(\Gamma_i) + E_{\text{lin}}(\Gamma_i) + E_{\text{per}}(\Gamma_i) + E_{\text{bnd}}(\Gamma_i) + \lambda_{\text{reg}}, \quad (4)$$

式中角速度约束项 E_{ang} 和线速度约束项 E_{lin} 用来奖励速度趋于稳定或变化平缓的轨迹,惩罚速度发生突变的轨迹; E_{per} 项约束轨迹以减少在不可能位置处开始或结束的短轨迹数量;边界条件 E_{bnd} 约束多项式在范围外达到极值; λ_{reg} 为一项恒定惩罚项,统一添加到所有轨迹中。

3 改进方法设计

3.1 外观约束项能量函数

在复杂环境中跟踪多个目标时,时常会发生身份标签互换问题,即在形成轨迹的过程中出现一段轨迹不属于同一目标的情况。如图 1 所示,若轨迹 Γ_1 与 Γ_2 中目标有相似的运动速度、边界条件等,在两目标相遇或相近时 DCEM 算法并不能有效区分目标。在数据关联的过程中,不同检测结果间的外观相似度可以用来衡量它们是否属于同一目标。据此,对轨迹选取不同位置检测结果计算相似度作为对轨迹的外观约束,以期在后续求解最小能量轨迹的过程中,避免选择有身份标签互换的轨迹。

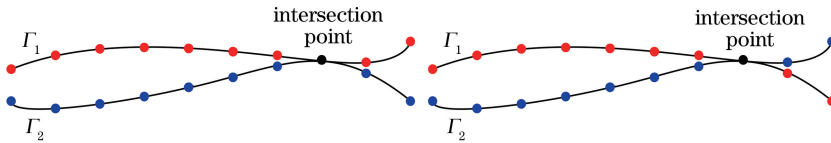


图 1 (a) 正常轨迹; (b) 身份标签互换轨迹

Fig. 1 (a) Normal trajectories; (b) trajectories with identity switch

为对轨迹建立外观约束,需提取轨迹目标图像的特征。在此对轨迹选取其中 4 帧目标图像进行特征计算。这 4 帧图像分别选取自轨迹的起始端、结束端和中间两点所对应图像的目标检测结果。如图 2 所示,红色虚线圈所指出的 4 个点分别对应起始点、中间点 1、中间点 2 和结束点。

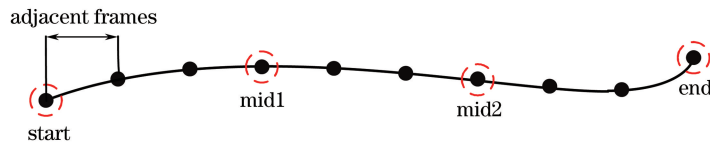


图 2 检测点选取示意图

Fig. 2 Detection point selection

为了对目标外观进行合适的表达,使其在复杂环境中具有良好的稳健性,采用多特征融合方式提取外观特征。考虑到常用几种局部特征的优缺点,在此选择目标的颜色特征、表象特征和纹理特征作为表征目标外观的几个局部特征。颜色特征选用 RGB 颜色直方图,将三通道的直方图串联,得到 96 维的目标颜色直方图向量 \mathbf{v}_{RGB} ;表象特征选用方向梯度直方图(HOG)特征,提取 32 维特征向量 \mathbf{v}_{HOG} ;纹理特征选用局部二值模式(LBP)特征,输入目标图像经过 LBP 算子运算后对图像进行直方图统计,得到 32 维 LBP 直方图特征向量 \mathbf{v}_{LBP} 。最终采用串联方式融合三个局部特征,得到 160 维的目标外观特征向量 $\mathbf{f} = \{\mathbf{v}_{\text{RGB}}, \mathbf{v}_{\text{HOG}}, \mathbf{v}_{\text{LBP}}\}$ 。

对一段轨迹 4 个点对应的目标检测结果利用上述方法进行外观特征向量计算,得到 4 个目标的特征向量分别为 $\mathbf{f}_{\text{start}}$ 、 \mathbf{f}_{mid1} 、 \mathbf{f}_{mid2} 、 \mathbf{f}_{end} 。为获得不同检测点之间外观特征的差别,采用常用的距离度量法——欧氏距离来度量。记 $\mathbf{f}_{\text{start}} = (x_{11}, x_{12}, \dots, x_{1n})$, $\mathbf{f}_{\text{mid1}} = (x_{21}, x_{22}, \dots, x_{2n})$, 其中 $n=160$ 。则 \mathbf{f}_{mid1} 与 $\mathbf{f}_{\text{start}}$ 之间的欧氏距离 d_{21} 为

$$d_{21} = \sqrt{\sum_{k=1}^n \left(\frac{x_{2k} - x_{1k}}{s_k} \right)^2}, \quad (5)$$

式中 s_k 为标准差。同理可以分别求得 \mathbf{f}_{mid2} 与 \mathbf{f}_{mid1} 、 \mathbf{f}_{end} 与 \mathbf{f}_{mid2} 和 $\mathbf{f}_{\text{start}}$ 与 \mathbf{f}_{end} 之间欧氏距离 d_{32} 、 d_{43} 和 d_{14} 。

定义外观约束能量函数为

$$E_{\text{app}}(\Gamma_i) = \lambda_{\text{app}} \lg(1 + d_f^2), \quad (6)$$

式中 λ_{app} 为外观约束项系数,即为外观约束能量在整个标签成本中所占权重,取值区间为 0 到 1; d_f 为 4 个外观特征向量欧氏距离的和,即 $d_f = d_{21} + d_{32} + d_{43} + d_{14}$ 。添加对轨迹的外观约束后的轨迹标签成本项为

$$E_{\text{traj}}(\Gamma_i) = E_{\text{app}}(\Gamma_i) + E_{\text{ang}}(\Gamma_i) + E_{\text{lin}}(\Gamma_i) + E_{\text{per}}(\Gamma_i) + E_{\text{bnd}}(\Gamma_i) + \lambda_{\text{reg}}. \quad (7)$$

外观约束能量函数奖励外观形变较少的轨迹,惩罚外观变化较大,尤其是发生突变的轨迹,以避免在跟踪中发生身份标签互换。通过实验得出当外观约束项系数 $\lambda_{\text{app}} = 0.8$ 时,能获得较为理想的跟踪结果。

3.2 后处理过程

对于存在遮挡的场景,由于可获取信息的缺失,时常会出现跟踪的轨迹分段而得到错误的轨迹。为此,对优化后的轨迹进行后处理,连接可能为同一目标的短轨迹。在数据关联方法中,不同轨迹的连接概率根据轨迹片间的运动、时间、外观相似度来计算。据此,采用时空邻域内短轨迹的外观和运动相似度来考量短轨迹的连接概率。

对优化后得到的轨迹 Γ_i ,在一定时空邻域内寻找其相邻轨迹 Γ_j ,两条轨迹有可能是同一轨迹的分段。为验证这两条轨迹段是否属于同一目标,对两者的相似度进行度量,则

$$S_{ij} = S_{\text{app}}(\Gamma_i, \Gamma_j) + S_{\text{motion}}(\Gamma_i, \Gamma_j), \quad (8)$$

式中 $S_{\text{app}}(\Gamma_i, \Gamma_j)$ 为外观相似度,利用(5)式对轨迹 i 结束端和轨迹 j 起始端对应的检测结果进行外观特征相似度计算得到。 $S_{\text{motion}}(\Gamma_i, \Gamma_j)$ 为运动相似度,测量运动的平滑度,采用文献[12]中的负对数高斯模型,即

$$S_{\text{motion}}(\Gamma_i, \Gamma_j) = -\ln G(p_i^{\text{end}} + v_j^{\text{end}} \Delta t_{ij} - p_j^{\text{start}}) - \ln G(p_j^{\text{start}} - v_j^{\text{start}} \Delta t_{ij} - p_i^{\text{end}}), \quad (9)$$

式中 p_i 和 v_i 分别表示轨迹 Γ_i 开始和结束部分的精确位置和线速度, Δt_{ij} 表示轨迹 i 的结束端和轨迹 j 的起始端之间的时间间隔。

在设定的时空阈值范围内若有不止一条相邻轨迹,分别对其计算相似度,时间阈值为 δ_t ,空间阈值为 δ_{sp} 。相似度 S_{ij} 值越小表明两条轨迹越相似,若 S_{ij} 小于设定的阈值 δ_s ,则认为两条轨迹是同一目标轨迹的分段,对其进行连接,并再次用三次 B-样条曲线拟合,得到最终的轨迹曲线。

3.3 算法步骤

在运动较为简单的场景中,如目标数量较少、无遮挡、无交叉运动等场景,仅考虑轨迹形态的约束就可以取得较好的跟踪结果,但实际环境的复杂性使得简单约束计算导致轨迹分段、身份标签互换等问题。为此设计如图 3 所示的算法,对检测结果生成的初始轨迹计算对应帧图像中目标的外观及运动等特征,生成轨迹约束能量;利用能量最小化方法迭代计算最小能量,获得能量最小时的所有目标轨迹;对所有轨迹根据相似度图像匹配理论融合分段轨迹,得到最终目标轨迹。

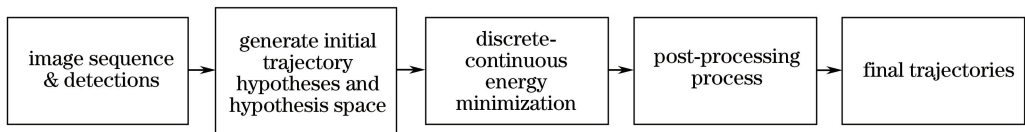


图 3 算法框图

Fig. 3 Diagram of the algorithm

具体步骤如下:

输入:图像序列 I 及其对应的检测结果 D 。

- 1) 读入图像序列 I 及检测结果 D ,生成初始轨迹 Γ 及轨迹假设空间^[11];
- 2) 选取轨迹检测点,提取对应外观特征向量,根据(6)式和(7)式分别计算外观约束能量和轨迹假设 Γ 中每条轨迹标签成本,之后计算(1)式的最小能量 E ,获得身份标签 f ;
- 3) 根据能量 E 修改身份标签 f ,利用新身份标签计算(1)式的最小能量 E ,获得新轨迹并更新轨迹假设空间;
- 4) 返回 2) 直至轨迹假设空间中所有假设轨迹均完成计算且并无更小能量 E ,得到身份标签 f 及轨迹 Γ ;
- 5) 根据(8)式计算时空邻域内相邻轨迹的相似度,若小于阈值则连接相邻轨迹并再次利用三次 B-样条

拟合,得到最终轨迹 Γ_f 。

输出:轨迹 Γ_f 。

4 实验结果及其分析

本文算法的实现基于 Windows 7 操作系统,采用 Matlab 2014a 为运行平台,计算机配置为 Intel(R) Core(TM) i5-6500 CPU @ 3.20 GHz 处理器,4.00 GB 内存,64 位操作系统。为了验证算法性能,在 Multiple Object Tracking Benchmark 上的公共数据集 2D MOT 15 的视频序列中进行实验,检测结果同样来自于该数据集。鉴于篇幅有限,这里只展示其中三段不同场景视频序列 PETS2009-S2L1、ADL-Rundle-6、TUD-Stadtmitte 的多目标跟踪数据结果,用不同颜色的矩形框表示对不同目标的跟踪。表 1 为三段视频的场说明,分别从拍摄视角、拍摄范围、目标运动方式等方面对测试视频场景进行说明。

表 1 测试视频场景说明

Table 1 Scene description of test video

Sequence name	PETS2009-S2L1	ADL-Rundle-6	TUD-Stadtmitte
Frame number	795	525	179
Resolution	768×576	1920×1080	640×480
View	High	Low	Low
Shooting range	Large	Small	Moderate
Target size	Small	Different sizes	Moderate
Target density	Moderate	Intensive	Moderate
Target movement mode	Non-linear motion	Linear & non-linear motion	Linear motion
Occlusion	Background occlusion & occlusion between targets	Occlusion between targets	Background occlusion
Tracking difficulty	High	High	Moderate
Time threshold	6	46	18
Space threshold	35	160	65

本文算法参数设置:外观约束项系数 λ_{app} 为外观约束能量在整个标签成本中所占权重,实验过程中从 0 到 1 每隔 0.1 取一次值进行实验,跟踪的准确度和精度随取值的增大而增大,在 0.8 时达到最大值,大于 0.8 时减小,因此取 $\lambda_{\text{app}}=0.8$ 。后处理过程中,由于实验视频拍摄角度与距离不同,时间阈值 δ_t 与空间阈值 δ_{sp} 的取值不同,具体值在表 1 中显示,时间阈值单位为帧,空间阈值单位为像素。相似度阈值 δ_s 在取 1.2 时,各项指标综合性能最佳。(7)式中其他参数与原算法保持一致。

图 4 为 PETS2009-S2L1 跟踪结果示意图,显示了原算法和改进后算法的跟踪效果。图 4(a)与(b)分别为原算法在第 124 帧和第 134 帧的跟踪结果,对比两帧可以看到,当粉红框黑衣行人与蓝框白衣行人相遇时,两者发生了身份标签互换,并且由于遮挡,路灯柱右面的目标没有被跟踪到。图 4(c)为改进后算法在第 134 帧的跟踪结果,图中可以明显看出身份标签没有互换,轨迹颜色方向未发生改变,并且被遮挡的目标被跟踪到。图 4(d)与(e)分别为原算法在第 696 帧和第 707 帧的跟踪结果,对比两帧可以看到,第 696 帧中左上角浅蓝色框的黑衣目标、图像右边缘深蓝色框黑衣目标以及路灯柱旁红衣目标都发生了轨迹分段。图 4(f)为改进后算法在第 707 帧的跟踪结果,对比可以发现上述三个目标未发生轨迹分段现象。由此可见,增加外观约束项和后处理过程后,轨迹分段和身份标签互换问题得到改善。

图 5 为 TUD-Stadtmitte 视频序列在三维空间中跟踪结果示意图。三维坐标系的 x 轴、 y 轴、 z 轴分别表示视频帧 x 方向图像坐标、 y 方向图像坐标、视频序列帧数,不同的颜色代表不同的轨迹。由于改进后算法使用了外观约束项,在能量最小化循环过程中对轨迹假设空间修正的结果可能与原算法不同,故改进后算法与原算法对不同目标的轨迹选择了不同的颜色。图 5(a)为原算法跟踪结果,可以明显地看出:轨迹 1 有轨迹分段现象,轨迹 2 与轨迹 3 有身份标签互换现象,轨迹 4 与轨迹 5 也有轨迹分段现象。图 5(b)为改进后算法的跟踪结果,可以看到,除了轨迹 4 仍有分段现象之外,其他问题都得到解决,且效果良好。

采用公共标准 CLEAR MOT^[12]和文献[14]中常用的评价指标来评估本文算法性能,即目标跟踪准确

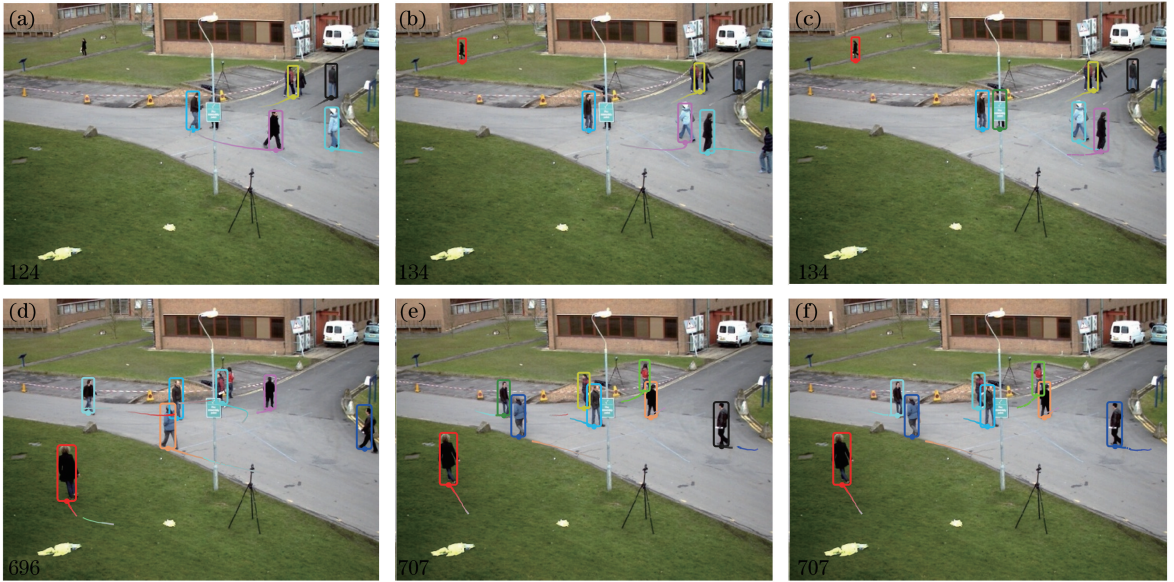


图 4 PETS2009-S2L1 序列跟踪结果。(a)~(c)身份标签互换；(d)~(f)轨迹分段

Fig. 4 Sequence PETS2009-S2L1 tracking results. (a)-(c) Identity switch; (d)-(f) trajectory segmentation

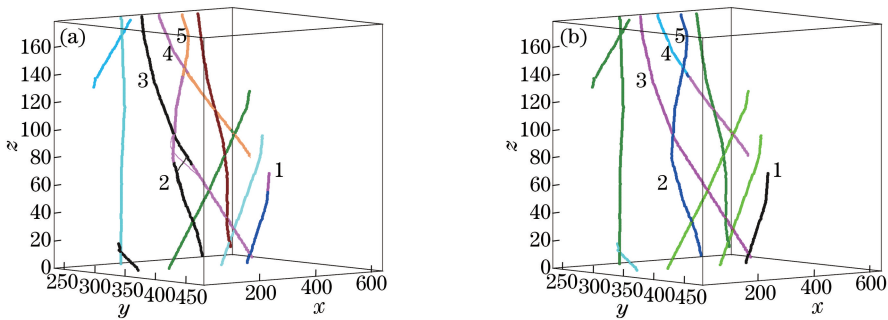


图 5 TUD-Stadtmitte 序列跟踪结果示意图。(a) DCEM 跟踪结果；(b)本文跟踪结果

Fig. 5 Sequence TUD-Stadtmitte tracking results. (a) Tracking results with DCEM; (b) tracking results with proposed method

度(MOTA)、目标跟踪的精确度(MOTP)、视频序列中真实轨迹数(GT)、在整个序列中超过 80% 视频帧被跟踪的轨迹数(MT)、在整个序列中低于 20% 视频帧被跟踪的轨迹数(ML)、视频序列目标真值所对应的轨迹分段的次数(FM)以及跟踪中目标身份标签转换次数(IDs)。表 2 为不同算法在不同测试视频序列上跟踪结果的对比,表中粗体表示最佳结果,其中 MOTA、MOTP 与 MT 值越大越好,GT 不变,ML、IDs、FM 越小越好。

表 2 中,三个对比文献分别为 Yoon 等^[15]提出的跟踪算法、Milan 等^[16]提出的连续能量最小化方法和文献[11]的原算法。其中文献[15]、[16]的数据为相关文献提供,文献[11]的数据是在图像平面实验所得。由表 2 可以看出,在三个测试序列中,改进后的算法在大部分指标上都有不同程度的提升。在 PETS2009-S2L1 视频序列中与文献[11]原算法对比,其跟踪准确度和跟踪精度都有所提升,同时身份标签转换和轨迹分段的情况大幅下降,但由于目标大多做非线性运动,这两种情况依然存在;与文献[15]和[16]相比,本文算法的 MOTA 和 MT 值低于文献[15]算法,表示跟踪准确度在一定程度上与其有差距,但差距不大,并且在轨迹分段和身份标签互换问题上本文算法最优。在 ADL-Rundle-6 视频序列中改进后的算法跟踪准确度有所提升但跟踪精度略有下降,同时身份标签转换及轨迹分段情况明显降低,尤其是身份标签转换问题得到明显改善。由于文献[15]和[16]未提供在该视频序列的实验结果,在此不进行对比。在 TUD-Stadtmitte 视频序列中,由于场景较为简单并且视频较短,原算法性能良好,但改进后算法总体效果仍有提高。与文献[15]和[16]相比,除跟踪精度不如文献[15]外,其他指标皆为最优。大量实验表明,本文算法可以有效解决

轨迹分段和身份标签互换问题,并且跟踪精度与跟踪准确度都有所提升。

表 2 实验结果对比

Table 2 Comparison of experimental results

Sequence name	Method	MOTA%	MOTP%	GT	MT	ML	IDs	FM
PETS2009-S2L1	Ref. [15]	66.6	57.5	19	16	0	40	38
	Ref. [16]	57.9	59.7	19	15	0	52	44
	Ref. [11]	59.8	74.1	19	13	0	60	72
	Proposed	62.8	75.6	19	15	0	32	36
ADL-Rundle-6	Ref. [11]	46.5	79.0	25	18	5	97	19
	Proposed	52.2	77.8	25	22	2	25	7
TUD-Stadtmitte	Ref. [15]	87.9	96.6	10	8	0	6	7
	Ref. [16]	71.1	65.5	10	7	0	4	3
	Ref. [11]	92.6	86.8	10	10	0	6	3
	Proposed	95.8	89.2	10	10	0	1	1

5 结 论

针对在复杂环境中离散-连续能量最小化多目标跟踪算法出现的轨迹分段与身份标签互换的问题,提出了一种改进方案。通过考虑跟踪目标的外观变化因素,在轨迹的标签约束中添加外观约束项改进原能量函数以及增加全局优化后的后处理过程,提高了算法性能。实验结果表明,改进后算法有良好的性能。但是对于目标外观相似性较高的场景,如以足球比赛中的各个运动员作为跟踪目标,则算法的稳健性还有待提升,这将是下一步研究工作的重点。

参 考 文 献

- [1] Ding Qi, Gu Guohua, Xu Fuyuan, *et al.* Moving target detection on moving camera with the presence of strong parallax [J]. *Laser & Optoelectronics Progress*, 2015, 52(9): 091501.
丁祺, 顾国华, 徐富元, 等. 强视差下的移动相机运动目标检测[J]. *激光与光电子学进展*, 2015, 52(9): 091501.
- [2] Mo Shaowen, Deng Xinpu, Wang Shuai, *et al.* Moving object detection algorithm based on improved visual background extractor[J]. *Acta Optica Sinica*, 2016, 36(6): 0615001.
莫邵文, 邓新蒲, 王帅, 等. 基于改进视觉背景提取的运动目标检测算法[J]. *光学学报*, 2016, 36(6): 0615001.
- [3] Wang Hongqiao, Cai Yanning, Fu Guangyuan, *et al.* Recognition and tracking of multiple slowly-moving ground targets based on image series[J]. *Laser & Optoelectronics Progress*, 2016, 53(5): 051501.
汪洪桥, 蔡艳宁, 付光远, 等. 基于图像序列的地面慢动多目标识别与跟踪[J]. *激光与光电子学进展*, 2016, 53(5): 051501.
- [4] Xiang Y, Alahi A, Savarese S. Learning to track: online multi-object tracking by decision making [C]. *International Conference on Computer Vision*, 2015: 4705-4713.
- [5] Riahi D, Bilodeau G A. Online multi-object tracking by detection based on generative appearance models[J]. *Computer Vision and Image Understanding*, 2016, 152: 88-102.
- [6] Zhao H, Qi S, Wang X. A joint appearance model of SRC and MFH for multi-objects tracking[J]. *Neurocomputing*, 2016, 213: 75-83.
- [7] Sanchez-Matilla R, Poiesi F, Cavallaro A. Online multi-target tracking with strong and weak detections[C]. *European Conference on Computer Vision*, 2016: 84-99.
- [8] Zhang L, Li Y, Nevatia R. Global data association for multi-object tracking using network flows[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2008: 1-8.
- [9] Pirsivash H, Ramanan D, Fowlkes C C. Globally-optimal greedy algorithms for tracking a variable number of objects [C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2011: 1201-1208.
- [10] Butt A A, Collins R T. Multi-target tracking by Lagrangian relaxation to min-cost network flow[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2013: 1846-1853.

- [11] Milan A, Schindler K, Roth S. Multi-target tracking by discrete-continuous energy minimization [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(10): 2054-2068.
- [12] Qin Z, Shelton C R. Improving multi-target tracking via social grouping[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2012: 1972-1978.
- [13] Stiefelwagen R, Bernardin K, Bowers R, *et al.* The CLEAR 2006 evaluation[C]. International Evaluation Workshop on Classification of Events, Activities and Relationships, 2006: 1-44.
- [14] Wu B, Nevatia R. Tracking of multiple, partially occluded humans based on static body part detection[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, 1: 951-958.
- [15] Yoon J H, Yang M H, Lim J, *et al.* Bayesian multi-object tracking using motion context from multiple objects[C]. IEEE Winter Conference on Applications of Computer Vision, 2015: 33-40.
- [16] Milan A, Roth S, Schindler K. Continuous energy minimization for multi-target tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(1): 58-72.