

基于目标检测及跟踪的视频摘要技术研究

田合雷¹ 丁胜² 于长伟¹ 周立²

(合肥工业大学仪器科学与光电工程学院 合肥 230009)¹ (合肥工业大学计算机与信息学院 合肥 230009)²

摘要 为了能够在不丢失有用信息的前提下对海量监控视频进行摘要,提出一种基于目标检测及跟踪的视频摘要技术。首先根据混合高斯模型对视频进行背景建模,获得监控视频的背景;然后通过背景减法对运动目标进行检测,得到前景运动目标;其次对检测到的运动目标采用分层关联的思想进行目标跟踪,得到运动目标的完整信息;最后将这些运动目标及视频背景重新组装成摘要视频。实验结果表明,该方法能够有效地对监控视频进行浓缩,形成的摘要视频能够完整地保存原有视频信息,减少了存储空间,降低了成本,方便了相关人员及时获取有用信息,提高了工作效率。

关键词 视频摘要,视频浓缩,目标检测,目标跟踪,背景建模,分层关联

中图分类号 TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.11.057

Research of Video Abstraction Based on Object Detection and Tracking

TIAN He-lei¹ DING Sheng² YU Chang-wei¹ ZHOU Li²

(School of Instrument Science and Opto-electronics Engineering, Hefei University of Technology, Hefei 230009, China)¹

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)²

Abstract In order to carry out a summary of massive surveillance video under the premise of not losing useful information, a video summarization technique based on object detection and tracking was proposed. Moving targets is detected by background subtraction in video by using Gaussian mixture model, and target tracking for detected target is done by using the idea of hierarchical correlation to get complete information of moving targets. Finally, the moving objects and video background are reassembled into the abstract video. The experimental results show that the method proposed in this paper can be used to concentrate the monitoring video. The summary of the video can completely preserve the original video, which reduces the storage space and cost. It is also convenient for relevant personnel to obtain useful information in time and improve work efficiency.

Keywords Video abstraction, Video condensation, Object detection, Object tracking, Background modeling, Hierarchical correlation

1 引言

视频摘要又称视频浓缩,是对视频内容的一种简单概括。通过保留感兴趣的运动目标的视频信息,剔除视频中冗余的内容,实现对一个监控视频中有用信息的筛选和浓缩。在不丢失视频信息内容的同时,尽量缩短视频的时间长度,得到视频的摘要。近年来,随着学校教育大众化和学校后勤社会化的推进,校园人员多元化,使得学校由原来封闭型的净土转变为开放性“小社会”。随着近年来校园暴力事件和突发事件的增加,传统的人力巡查已不能满足校园安全管理的需求,越来越多的学校开始考虑通过校园网络视频监控系统的实施来建设平安校园。然而校园监控视频录像存在存储数据量大、存储时间长等特点,通过原始录像寻找线索进而获取证据的传统做法要耗费大量人力、物力以及时间,效率极其低下,以至于错过最佳破案时机。因此在视频监控系统中,为了达到快速浏览、及时获取有用信息、锁定可疑对象的目的,对原始视频进行浓缩变得尤为重要。

目前国内外有不少学者在这方面展开研究,并取得了一定的成果。文献[1-3]基于提取关键帧的方法来对视频进行摘要浓缩,即将视频分成场景、镜头和关键帧几个层次,最后保留关键帧;关键帧的播放是跳跃、不连续的,并且视频内容复杂时将导致关键帧提取不合理,可能会丢失关键信息,同时由于形成摘要视频播放时不连贯,人的视觉感觉会比较难受,用户体验比较差。文献[4-6]基于对象的视频摘要,通过目标跟踪获取目标对象的完整信息。首先将背景和活动目标分离,然后将目标回贴到背景上,用空间换取时间,不同时刻的目标出现在同一画面中,达到压缩视频的目的。但是现有技术还存在一些问题,比如,若背景上叠加的对象过多则会干扰人工排查;如果目标检测的准确性、目标跟踪的连续性不高,就会导致摘要重复、混乱等问题。

针对以上问题,本文提出基于目标检测及跟踪的视频摘要技术,该技术能够准确获得运动目标;再采用分层关联的目标跟踪获取运动目标的完整信息;然后按照浓缩比形成摘要视频。整个技术过程如图 1 所示。

到稿日期:2015-10-23 返修日期:2016-01-05 本文受国家自然科学基金(61371155),安徽省科技攻关基金项目(1301b042023)资助。

田合雷(1976-),男,硕士,讲师,主要研究方向为人工智能与模式识别、传感器技术, E-mail: thl@hfut.edu.cn; 丁胜(1979-),男,硕士,讲师,主要研究方向为智能信息处理、数字图像分析与处理; 于长伟(1979-),男,硕士,讲师,主要研究方向为精密测试技术; 周立(1979-),男,工程师,主要研究方向为信息系统与智能决策。

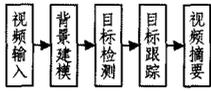


图1 视频摘要过程

本文第2节介绍运动目标检测;第3节介绍目标跟踪;第4节介绍视频摘要浓缩;第5节介绍实验结果;最后总结全文。

2 运动目标检测

在目标检测算法^[7-9]中,基于统计的方法实现背景建模和目标检测的效果较好。该方法利用统计像素点的值来表征背景,从而建立背景模型。在不同的监控场景下,其背景模型具有不同的特性。有文献将该方法单独分为一类,称为统计背景模型法,但该算法也是通过对比当前帧图像与背景帧图像的差异值来实现目标检测,因此本文将将其归入背景减法这一类。

Stauffer等人提出了经典的自适应混合高斯背景模型^[10,11]。它是一种基于像素的背景建模方法,利用像素在时间域的分布情况来构建像素点的颜色分布模型,从而建立背景模型。混合高斯背景模型是有限个高斯函数的加权和,它能够描述像素点的多峰状态,适合对由光照等引起的背景缓慢变化以及树木摇摆等复杂背景进行准确建模。具体流程如下。

对视频序列中的每一帧中的所有像素进行建模,设某位置为 (x_0, y_0) 的像素点在一段时间段内的观测值为:

$$\{X_1, \dots, X_t\} = \{I(x_0, y_0, i); 1 \leq i \leq t\} \quad (1)$$

使用 K 个高斯分布的拟合来描述像素点 (x_0, y_0) 的观测值,可以得到当前点颜色值的概率为:

$$f(X_t) = \sum_{i=1}^k \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2)$$

$$\eta(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_{i,t}|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T \Sigma_{i,t}^{-1} (X_t - \mu_{i,t})} \quad (3)$$

其中, K 为高斯函数的个数,取值为3-5; $\omega_{i,t}$ 为灰度值符合第 i 个高斯分布的像素点个数占所有像素点个数的比例(即模型函数的权重系数); $\mu_{i,t}$ 为第 i 个高斯函数的均值; $\Sigma_{i,t}$ 为第 i 个高斯函数的协方差矩阵; n 为任一像素观测向量 X_t 的维数。

为了计算方便,一般假定向量 X_t 的元素(即R,G,B3个分量)相互独立且具有相同的方差,则有:

$$\Sigma_{i,t} = \sigma_i^2 I \quad (4)$$

现在已经建立被观察像素点 (x_0, y_0) 颜色值的高斯混合模型。对于输入图像中任一像素点 (x_0, y_0, t) ,首先确定与它最近接的高斯分布,将其与已经存在的 K 个高斯分布逐个进行匹配,如果像素点 (x_0, y_0, t) 的亮度值与某个高斯分布的均值的差值满足条件,则认为该像素点和这个高斯分布“匹配”,即满足:

$$|(X_t - \mu_{i,t-1})| < TH * \sigma_{i,t-1} \quad (5)$$

其中, TH 为设定偏差的门限值,一般取2.5; $\mu_{i,t-1}$ 为第 i 个高斯分布在 $t-1$ 时刻的均值; $\sigma_{i,t-1}$ 为第 i 个高斯分布在 $t-1$ 时刻的偏差。对于满足上述条件的高斯分布,按如下公式更新该模型的各个参数。

$$\omega_{i,t} = (1 - \alpha_i) \omega_{i,t-1} + \alpha_i \quad (6)$$

$$\mu_{i,t} = (1 - \beta_i) \mu_{i,t-1} + \beta_i X_t \quad (7)$$

$$\sigma_{i,t}^2 = (1 - \beta_i) \sigma_{i,t-1}^2 + \beta_i (X_t - \mu_{i,t})^T (X_t - \mu_{i,t}) \quad (8)$$

其中, α_i 为模型的学习速率, $0 \leq \alpha_i \leq 1$; β_i 为模型参数的学习速率,可近似用 $\alpha_i / \omega_{i,t}$ 表示。

学习速率 α_i 值的大小对于背景建模非常重要。若 α_i 取

值过小,则模型更新速度缓慢,会把某些已经不是前景的物体判断为运动目标,造成目标的误检;若 α_i 取值过大,则模型更新速度过快,容易把运动目标更新到背景中,造成目标漏检。

若当前像素的灰度值没有与任何一个高斯模型匹配,则使用一个新的高斯模型替代当前高斯模型中权重最小的高斯分布,然后可以再通过计算任一高斯分布 $\omega_{i,t} / \sigma_{i,t-1}$ 的值来进行判断,若该值越大,则该像素点是背景点的可能性就越大。

混合高斯模型方法能够表征像素点灰度值分布的多模态,在背景建模过程中允许有行人、车辆等运动的前景目标。该方法可应用于复杂场景中光照缓慢变化以及存在重复运动目标的背景的背景的准确建模,尤其适用于有光线缓慢变化和天气变化小而目标速度运动较快的目标检测。

3 目标跟踪

目标跟踪一直是计算机视觉领域的研究热点之一,是视频分析与智能视频监控系统中一项基础而又关键的任务。目标跟踪的目的就是将观测目标在场景中的运动轨迹标记出来,一种成功的跟踪算法能够始终在视频序列的各个图像帧中检测并定位同一目标。本文采用分层关联^[12-14]的方法实现目标跟踪,首先通过相邻帧的匹配得到初始轨迹片段,然后对轨迹片段进行进一步优化得到目标运动轨迹。

3.1 初始轨迹片段

首先对相邻帧的检测点计算出彼此的关系亲密度,然后通过阈值策略判断某两个检测点是否属于同一个目标,最后对满足阈值要求的检测点建立关联。

3.1.1 关系亲密度的计算

假设通过运动目标检测得到的检测点集合为 $D = \{d_i\}$, $d_i = (x_i, y_i, s_i, t_i, a_i)$ 。其中 (x_i, y_i) 表示检测点 d_i 的位置, s_i 表示大小, t_i 表示该检测点所在的当前帧号, a_i 表示检测点的颜色直方图。存在某两个检测点 d_i, d_j ,它们的关系亲密度由下式得到。

$$P(d_i, d_j) = \begin{cases} A_{ij}^{pos} A_{ij}^{size} A_{ij}^{app}, & \text{if } t_j - t_i = 1 \\ 0, & \text{其他} \end{cases} \quad (9)$$

位置、面积与外观特征相似性计算可分别表示为:

$$A_{ij}^{pos} = \lambda_{pos} \exp\left[-\frac{(x_i - x_j)^2}{\sigma_x^2}\right] \exp\left[-\frac{(y_i - y_j)^2}{\sigma_y^2}\right] \quad (10)$$

$$A_{ij}^{size} = \lambda_{size} \exp\left[-\frac{(s_i - s_j)^2}{\sigma_s^2}\right] \quad (11)$$

$$A_{ij}^{app} = D(a_i, a_j) \quad (12)$$

其中, $D(\cdot)$ 是两个检测点颜色直方图的巴氏距离; $\lambda_{pos}, \lambda_{size}$ 是权重系数,用来平衡三者对总体结果的影响; σ_x, σ_y 分别是轨迹片段的 x 轴坐标和 y 轴坐标的方差; σ_s 是当前轨迹片段面积的方差。

3.1.2 双阈值判定

当一个检测点与相邻帧中的多个检测点都有很高的关系亲密度时,若只用一个阈值来判定,往往结果并不可靠。本文为了避免这种不安全的关联,采用双阈值的策略:

$$P(d_i, d_j) > \theta_1 \\ \min[P(d_i, d_j) - P(d_k, d_j), P(d_i, d_j) - P(d_i, d_k)] > \theta_2 \quad (13)$$

$$\text{s. t. } \theta_1 > \theta_2 > 0 \text{ and } d_k \neq \{d_i, d_j\}$$

阈值 θ_1 表示两个检测点必须有很高的亲密度,同时阈值 θ_2 是为了保证这两个检测点亲密度高于其他任何两个检测点的亲密度。只有满足双阈值的两个检测点才被关联。

3.1.3 轨迹片段的起止帧的确定

显然,检测的第一帧出现的目标都应是轨迹片段的起始,

然后在下一帧中找相似的目标。如果下一帧没有相似的目标,就在后一帧继续寻找。最大检索帧差数为 5,即 5 帧之内没有找到相似的目标,就认为这个轨迹片段结束的位置已经找到。在视频序列中出现的目标是不是一个新的轨迹片段的开始,取决于这个目标是否被前面的目标匹配上,如果没有,就是新的轨迹片段的开始。

3.2 轨迹片段优化

轨迹片段优化也叫轨迹片段关联,把上一节得到的轨迹片段进行进一步关联优化得到目标的整个运动轨迹。设轨迹片段的集合 $T = \{T_i\}$, $T_i = (t_i^s, t_i^e, p_i^s, p_i^e, v_i^s, v_i^e, s_i, a_i)$ 。其中 $t_i^s, t_i^e, p_i^s, p_i^e, v_i^s, v_i^e, s_i, a_i$ 分别表示轨迹片段的起始帧、结束帧、起始位置、结束位置、起始速度、结束速度、外观大小、外观特征。轨迹片段的外观特征 a_i 是整个轨迹片段中所有检测点外观特征的中值,外观大小 s_i 是整个轨迹片段内所有检测点的平均大小。存在两个轨迹片段 T_i, T_j ,两者的关系亲密度可由下式计算:

$$P(T_i, T_j) = P_s(T_i, T_j)P_{size}(T_i, T_j)P_{app}(T_i, T_j) \quad (14)$$

其中,面积和外观相似性 $P_{size}(T_i, T_j), P_{app}(T_i, T_j)$ 的计算与 $A_{ij}^{size}, A_{ij}^{app}$ 的计算类似。 $P_s(T_i, T_j)$ 表示轨迹片段之间的时空相似性。

$$P_s(T_i, T_j) = \lambda_s [e(T_i, T_j) + e(T_j, T_i)] \quad (15)$$

$$e(T_i, T_j) = \|q_i - p_j\|_2 \quad (16)$$

$$q_i = p_i^s + v_i^s(t_j^s - t_i^s) \quad (17)$$

$$e(T_j, T_i) = \|q_j - p_i^e\|_2 \quad (18)$$

$$q_j = p_j^e + (-v_j^e(t_j^e - t_i^e)) \quad (19)$$

其中, λ_s 表示关系权重。式(16)、式(18)分别表示轨迹片段的前向匹配误差和后向匹配误差。同样采用类似于式(13)的双阈值的策略来判定两个轨迹片段是否被关联。

4 视频摘要浓缩

经过背景建模,得到一张没有运动物体的清晰背景图像。然后进行运动目标检测及目标跟踪,得到所有运动目标的信息,包括位置、起始时刻以及目标图片,将这些视频有用信息从原来的视频中提取出来。接下来的工作是把这些有用信息结合视频背景重新组装成新的视频,得到视频摘要,达到视频浓缩的目的。具体步骤如下:

- (1)通过背景建模得到背景图像帧;
- (2)对经过目标跟踪得到的运动目标,按照时间顺序在背景图像上进行放置;
- (3)对不同时刻的目标在同一位置有重叠的部分进行半透明处理;
- (4)当一帧图像中的目标数大于指定的数目时,不再继续放置目标;
- (5)将步骤(4)中的图像帧保存到视频序列中,复制新的背景图像帧,继续放置目标直至放完;
- (6)当所有的图像帧写入到视频流中后,就形成了最终的视频摘要。

5 实验结果

本文提出的基于目标检测及跟踪的视频摘要算法使用 C++ 和 OpenCV2.4.9 编程,并在处理器为 Interl(R)Core(TM)i7 4.00GHz,内存 8GB,Windows7 操作系统的 PC 上完成了实验测试。本文的测试视频采用校园监控系统中实际采集到的视频序列。实验结果如图 2、图 3 所示。



图 2 目标检测与跟踪

图 2(a)是原始视频输入。图 2(b)是采用混合高斯背景建模并经过中值滤波得到的目标检测的结果。在本文实验中学习速率 α_i 取值为 0.001。图 2(c)中的黑色框是通过目标检测得到的运动目标。图 2(d)是运动目标的跟踪结果,不同的目标使用不同的颜色标识。图 3(a)是不同时刻在同一场景出现的 4 个不同的目标,图 3(b)把图 3(a)中 4 个不同时刻出现的目标贴在同一帧中,形成视频摘要。实验结果表明,采用本文提出的方法所得目标检测结果比较准确,能够对目标进行持续的跟踪,形成的摘要视频能够有效地保存原有视频信息,包括时间信息、空间信息,播放流畅,方便相关人员查看。



图 3 生成摘要视频

结束语 本文提出的基于目标检测及跟踪的视频摘要算法首先通过混合高斯模型背景建模并利用背景减法进行运动目标的检测,然后通过分层关联对运动目标进行跟踪,提取视频的有用信息,最后根据运动目标的有用信息剔除视频的冗余信息,并结合背景图片融合生成摘要视频。实验结果表明,本文算法能够在不丢失视频信息的情况下对视频进行有效压缩。但是当视频场景内情况比较复杂时,对目标检测及跟踪会有一些影响,从而会出现间断及跟踪中断的现象,避免这种现象将是我们的下一步的工作重点。同时我们进一步的研究方向是针对摘要视频进行特定行人和车辆检索。

参考文献

- [1] Kumar N S, Shobha G, Balaji S. Key frame extraction algorithm for video abstraction applications in underwater videos[C]//Underwater Technology(UT), 2015. IEEE, 2015: 1-5

(下转第 312 页)

参考文献

- [1] Ling kai, Wu Jun-hui, Xu Li-hong. A Survey on Color Image Segmentation Techniques[J]. Journal of Image and Graphics, 2005,10(1):1-10(in Chinese)
林开颜,吴军辉,徐立鸿.彩色图像分割方法综述[J].中国图象图形学报,2005,10(1):1-10
- [2] Rother C, Kolmogorov V, Blake A. "GrabCut"—interactive foreground extraction using iterated graph cuts[J]. Acm Transactions on Graphics, 2004, 23(3): 309-314
- [3] Ning J, Zhang L, Zhang D, et al. Interactive image segmentation by maximal similarity based region merging[J]. Pattern Recognition, 2010, 43(2): 445-456
- [4] Liu Song-tao, Yin Fu-liang. The basic principle and its new advances of image segmentation methods based on graph cuts[J]. Acta Automatica Sinica, 2012, 38(6): 911-922(in Chinese)
刘松涛,殷福亮.基于图割的图像分割方法及其新进展[J].自动化学报,2012,38(6):911-922
- [5] Rother C, Minka T, Blake A, et al. Cosegmentation of Image Pairs by Histogram Matching-Incorporating a Global Constraint into MRFs[C]// 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2006: 993-1000
- [6] Hochbaum D S, Singh V. An efficient algorithm for Co-segmentation[C]// IEEE International Conference on Computer Vision. 2009: 269-276
- [7] Bach F, Ponce J. Discriminative clustering for image co-segmentation[C]// IEEE Conference on Computer Vision & Pattern Recognition. 2010: 1943-1950
- [8] Peng J, Singh V, Mukherjee L. Scale invariant cosegmentation for image groups[C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012: 1881-1888
- [9] Collins M D, Xu J, Grady L, et al. Random walks based multi-image segmentation: Quasiconvexity results and GPU-based solutions[C]// 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2012: 1656-1663
- [10] Meng F, Li H, Liu G, et al. Image Cosegmentation by Incorporating Color Reward Strategy and Active Contour Model[J]. IEEE Transactions on Cybernetics, 2012, 43(2): 725-737
- [11] Batra D, Kowdle A, Parikh D, et al. iCoseg: Interactive co-segmentation with intelligent scribble guidance[C]// 2013 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2013: 3169-3176
- [12] Gulshan V, Rother C, Criminisi A, et al. Geodesic Star Convexity For Interactive Image Segmentation[C]// 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2010: 3129-3136
- [13] Wang Chun-yao, Chen Jun-zhou, Li Wei. Superpixel segmentation algorithms review[J]. Application Research of Computers, 2014, 31(1): 6-12(in Chinese)
王春瑶,陈俊周,李炜.超像素分割算法研究综述[J].计算机应用研究,2014,31(1):6-12
- [14] Arbelaz, Pablo, Maire M, et al. Contour Detection and Hierarchical Image Segmentation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(5): 898-916
-
- (上接第 299 页)
- [2] Song G H, Ji Q G, Lu Z M, et al. A novel video abstraction method based on fast clustering of the regions of interest in key frames[J]. AEU-International Journal of Electronics and Communications, 2014, 68(8): 783-794
- [3] Wang Ya-pei, Li Ren-wang, Liu Xiang. Extraction Method of Surveillance Video Synopsis Combines Objects and Keyframes [J]. Industrial Control Computer, 2015(3): 11-13(in Chinese)
王亚沛,李仁旺,刘翔.对象和关键帧相结合的监控视频摘要提取方法[J].工业控制计算机,2015(3):11-13
- [4] Zhang P, Zhuo T, Zhang Y, et al. Real-time Tracking-by-Learning with High-order Regularization Fusion for Big Video Abstraction[J]. Signal Processing, 2015, 124(c): 246-258
- [5] Liu Shou-da. Video Synopsis Based on Multi-target Tracking and Trajectory Combinatorial Optimization[D]. Xiamen: Xiamen University, 2014(in Chinese)
刘守达.基于多目标跟踪及轨迹组合优化的视频摘要[D].厦门:厦门大学,2014
- [6] Han Jan-kang. Moving Area Detection and Tracking Based Video Condensation[D]. Beijing: Beijing University of Posts and Telecommunications, 2012(in Chinese)
韩建康.基于运动检测及跟踪的视频浓缩方法研究[D].北京:北京邮电大学,2012
- [7] Felzenszwalb P, Girshick R, McAllester D, et al. Object detection with discriminatively trained part based models[J]. Pattern Analysis and Machine Intelligence (PAMI), 2010, 32(9): 1627-1645
- [8] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C]// Computer Vision and Pattern Recognition (CVPR). San Diego, CA, 2005, 1: 886-893
- [9] Wu B, Nevatia R. Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors[J]. International Journal of Computer Vision, 2007, 75(2): 247-266
- [10] Song Xue-hua, Chen Yu, Geng Jian-feng. Moving object detection based in improved Gaussian mixture background model [J]. Computer Engineering and Design, 2010, 31(21): 4646-4649 (in Chinese)
宋雪桦,陈瑜,耿剑锋.基于改进的混合高斯背景模型的运动目标检测[J].计算机工程与设计,2010,31(21):4646-4649
- [11] Zivkovic Z. Improved Adaptive Gaussian Mixture Model for Background Subtraction[C]// Proceedings of the International Conference on Recognition Pattern, Amsterdam University. Netherlands, 2004, 2: 23-26
- [12] Huang C, Li Y, Nevatia R. Multiple target tracking by learning-based hierarchical association of detection responses[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(4): 898-910
- [13] Wang Jiang-feng. Researches on Object Tracking and Event Detection Based on Tracklet Association[D]. Changsha: National University of Defense Technology, 2011(in Chinese)
王江峰.基于轨迹片段关联的目标跟踪与事件检测方法研究[D].长沙:国防科学技术大学,2011
- [14] Ristani E, Tomasi C. Tracking Multiple People Online and in Real Time[M]// Computer Vision—ACCV 2014. Springer International Publishing, 2015: 444-459