

跨摄像头目标跟踪综述

王先彬, 严明

(四川大学计算机学院, 成都 610065)

摘要:

跨摄像头目标跟踪是在多个摄像头之间跟踪具有时间和空间离散性的运动目标,在智能交通、公共安全等领域具有重要应用价值。针对跨摄像头跟踪任务的复杂性,先简要介绍跨摄像头目标跟踪的基本框架,然后详细分析跨摄像头目标跟踪的主要挑战和研究现状,重点介绍摄像头关联模型及图像融合算法,并探讨跨摄像头目标跟踪的发展趋势。

关键词:

跨摄像头; 目标跟踪; 摄像头关联模型; 图像融合

基金项目:

四川省科技厅苗子工程(No.0030405301461)

0 引言

跨摄像头目标跟踪是当前视频目标跟踪领域的研究热点,其主要任务是跟踪目标在多个摄像头之间的运动,是智能视频分析的基础。与单摄像头下的目标跟踪相比,它能够提供更对运动目标跨视域、大范围、长时间、鲁棒性更好的跟踪效果,在智能交通、公共安全等领域有重要应用价值。

跨摄像头目标跟踪研究是困难的,因为它不但继承了单摄像头目标跟踪的复杂性,还需要建立摄像头关联模型,融合来自各个摄像头的图像信息。本文先简要介绍跨摄像头目标跟踪的基本框架,再详细阐述跨摄像头目标跟踪的主要挑战和研究现状,重点介绍了摄像头关联模型及图像融合算法,分析其优缺点,最后对跨摄像头跟踪的未来进行展望。

1 跨摄像头目标跟踪的基本框架

跨摄像头目标跟踪是视频目标跟踪的子领域,由于运动目标类型、跟踪应用场景的不同,跨摄像头目标跟踪系统的实现也不同。例如,当用于楼内安保监控时,由于运动目标是人、摄像头视域狭窄及目标运动路

径受到物理限制,可以通过人工标定和校准方式来建立摄像头间的协同跟踪模型。而当目标是车辆,且要求在城市道路实现不间断跟踪时,跨摄像头跟踪就变得非常复杂。通过分析不同系统的实现,可以抽象出一般性的跨摄像头目标跟踪框架,该框架可以分为系统初始化模块、单摄像头跟踪模块和跨摄像头目标定位模块,如图1所示。

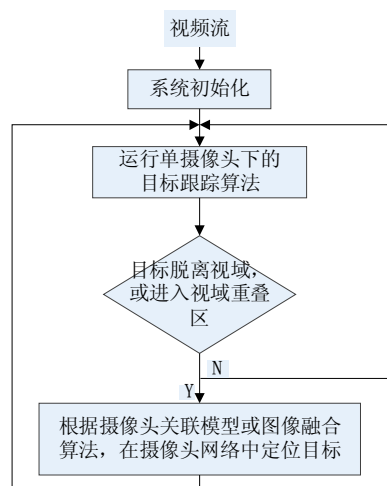


图1 跨摄像头目标跟踪基本框架

各模块的功能如下:

(1)系统初始化模块:接入各摄像头的输入视频,人工或通过目标检测算法选择被跟踪目标。

(2)单摄像头跟踪模块:包括建立运动模型、提取目标特征、选择外观模型、确定目标位置、更新外观模型等步骤。运动模型根据目标在上一帧中的位置,按照一定的规则产生目标在当前帧中的候选位置,常见的运动模型有粒子滤波、滑动窗口和均值漂移。提取目标特征有两种方法,一是通过人工设计的算法获得目标的点、几何形状、轮廓等形状特征,或灰度、颜色、纹理、梯度等外观特征;二是利用机器学习特别是深度学习自动获取边缘、颜色等底层特征和人脸、轮毂等高层抽象特征。外观模型分为生产模型和判别模型。生产模型需要在跟踪前对目标外观(如形状和纹理)建模,然后在运动模型基于上一帧产生的候选位置框中,选择与此模型最匹配的位置框作为目标在当前帧中的位置。判别模型也称为基于目标检测的模型,它利用一个在线分类器将目标与背景分离,从而确定目标的位置。更新外观模型主要是更新模板(对生产模型)或利用新样本对分类器进行增量式在线训练(对判别模型)。

(3)跨摄像头目标定位模块:建立和应用摄像头关联模型或图像融合算法,在相邻摄像头中确定目标的位置。该模块是实现跨摄像头目标跟踪系统的关键,它包括两个部分:一是线下建立摄像头关联模型或设计图像融合算法;二是线上更新该模型或算法。

2 跨摄像头目标跟踪的主要挑战

跨摄像头目标跟踪作为视频目标跟踪领域的一个研究分支,继承了传统的单摄像头目标跟踪研究的复杂性。这些复杂性包括但不限于以下几点:(1)场景光照变化;(2)目标被部分甚至完全遮挡;(3)目标运动复杂;(4)目标有刚性与非刚性之分;(5)视频图像采集时损失深度信息;(6)要求实时跟踪。跨摄像头跟踪实现大范围连续跟踪优点的同时,也引入了新的挑战。一方面,由于监控角度、摄像头景深以及像素差异等原因,同一个目标在不同摄像头拍摄的视频图像中会出现不同大小、视角、清晰度、运动速度等问题,需要进行图像融合。另一方面,目标离开一个摄像头视域后进入另一个摄像头视域是涉及到摄像头间的协同关联问题,是实现跨摄像头目标跟踪的关键。本文主要介绍跨

摄像头时目标跟踪的研究情况,而单摄像头目标跟踪算法是一个独立研究领域,在后文中不再涉及,感兴趣的读者可参考相关文献[1-3]。

3 跨摄像头目标跟踪研究现状

当前跨摄像头目标跟踪研究主要有两个分支:一是非重叠摄像头跟踪,是指两个摄像头之间的视域没有交叉,目标离开一个摄像头视域运动到另一个摄像头视域的过程中存在监控盲区,这种情况下摄像头关联模型是研究重点;另一个是摄像头视域存在重叠时的目标跟踪,即同一个目标同一时刻出现在多个摄像头视域中,这种情况下图像融合是研究重点。

3.1 非重叠摄像头跟踪方法及性能分析

(1)基于全局检测的摄像头关联方法

由于计算能力的极大增强以及摄像头采集的视频图像分辨率的提高,基于检测的跟踪成为近年来单摄像头目标跟踪领域的热门研究方向^[4]。在跨摄像头跟踪领域,也有不少学者提出了基于检测的跟踪方法^[5]。

Yong-Feng HUANG 等^[6]提出了一种基于 TLD 算法的非重叠摄像头跟踪系统。该系统的检测模块在目标离开某摄像头视域后,会全局扫描一定范围内摄像头的视频帧,然后通过模板匹配的方法定位目标。在单摄像头下跟踪目标时,作者用 MeanShift 和粒子滤波跟踪方法替换了 TLD 算法中基于特征点的光流跟踪方法,增强了在光照剧烈变化等复杂环境下系统的鲁棒性。但由于目标检测模块只采用简单的全局搜索方法,系统需要很强的计算能力,实时性比较差。

Jorge C. 等^[7]研究在隧道中跨摄像头跟踪车辆时,用 AdaBoost 分类器来加速摄像头间视频帧的全局搜索过程,即每个候选位置框同时被几个特征分类器评估,每个分类器都会剔除一半的候选位置框。这样大部分负样本在早期就会被剔除,从而加快了检测速度。剩下的正样本进入下一轮评估,最后得到的最优位置框就作为下一个摄像头中目标的位置。

李志华等^[8]提出用 SIFT 特征匹配多个摄像头中的目标的方法实现跨摄像头跟踪。当目标离开某摄像头视域后,中心服务器利用摄像头的相邻性检测该摄像头的相邻摄像头组中新出现的目标,并利用此新目标分割图像的 SIFT 特征与数据库中其他摄像头中出现过的目标进行特征匹配,若匹配成功,则将原摄像头下

目标的编号赋予新的摄像头下的目标。SIFT 特征对图像尺度、旋转、仿射变换以及视角方向变化等具有高度不变性,使得该方法能够达到较好的匹配跟踪效果。但是这种方法需要提前标定每个摄像头的相邻组,而且一旦目标离开一个摄像头视域后,中心服务区就需要全局搜索该摄像头的相邻组,需要网络和服务器能够承受较大负荷。

陈伟宏等^[9]通过结合摄像头间的目标外形变化(用特征距离 D 度量)和运动路径模型(用空间和速度特征度量),融合多个摄像头信息,减少了跨摄像头系统建立亮度变换函数(BTF)^[10]的学习时间,使得该算法能够快速识别出穿过摄像头的目标,而且对场景光照改变的不敏感。另外,该系统也不需要提前校准摄像头,从而增强了系统的扩展性。

(2) 基于机器学习的摄像头关联模型

行人从一个摄像头到另一个摄像头的过程中不穿过其他摄像头时,这两个摄像头称为直接连接(directly-connected)。直接连接的摄像头之间的路径包含一个进入区和一个退出区。训练数据是两组由进入区观测实例集(X)和退出区观测实例集(Y)组成,每个进入/退出观测实例包含行人的颜色、纹理特征以及进入/退出该视域(FOV)的时间。如果在摄像头数量较少且已知摄像头间拓扑图的情况下,人工标定直接连接的摄像头对以及与其关联的进入区/退出区是比较容易的。此时直接采用监督学习方法就能够获得摄像头关联模型。然而,人工标定训练数据不仅是一件费时费力的事,而且随着摄像头网络规模的增长变得越来越不可行。同时,每增加一个摄像头都需要重新标定训练数据,扩展性比较差。这种方式在商场、办公大楼等摄像头数量稳定的内部监控场合比较适用。

Chun-Te Chu^[11]提出利用一种非监督学习框架构建的摄像头关联模型来进行跨摄像头行人跟踪。文中将两个观测实例集 X 和 Y 之间的关联性表示为关联矩阵 $P_{(N_1+1) \times (N_2+1)}$,其中 N_1 和 N_2 分别是数据集 X 和 Y 中的样本个数。与监督学习方法中 P 为人工标定的二值矩阵不同,这里求解关联矩阵 P 是一个最优化问题,即给定最新摄像头关联模型 M 后由最小化代价函数簇 $J(\cdot)$ 得到,见式(1)。

$$\hat{P} = \operatorname{argmin}_P J(P) \quad (1)$$

$$s.t. \quad P_{ij} \in \{0, 1\} \quad \forall i \leq N_1+1, j \leq N_2+1 \quad (2)$$

$$\sum_{i=1}^{N_1+1} P_{ij} = 1 \quad \forall j \leq N_2, \sum_{j=1}^{N_2+1} P_{ij} = 1 \quad \forall i \leq N_1 \quad (3)$$

其中 P_{ij} 是连续实数,表示第 i 个退出观测实例与第 j 个进入观测实例的匹配程度; $J(\cdot)$ 中的每个代价函数表示某个特征在进入观测实例与退出观测实例之间的距离。求得 P 后,直接连接的摄像头之间的转移时间分布、亮度转移函数、区域匹配矩阵、区域匹配权值和特征融合权值就可由最新的 P 得到,从而获得更新的摄像头关联模型 M 。通过不断迭代更新,最后收敛得到稳定的关联矩阵 P 和摄像头关联模型 M 。整个流程分为训练阶段和测试阶段,见图 2。

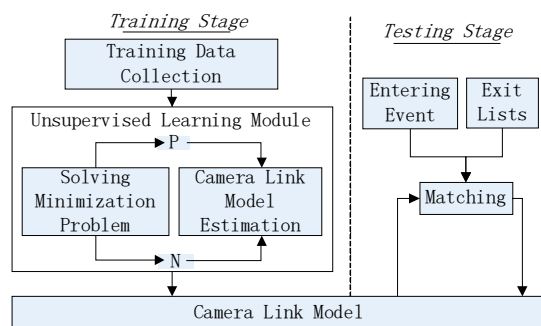


图2 摄像头关联模型构建图

该模型在有新摄像头或新的目标特征加入时有较好的泛化能力,但缺点是需要事先知道一下先验知识:①摄像头网络拓扑图;②哪些摄像头之间是直接连接(即存在关联模型);③一对直接连接的摄像头之间的关联意味着有且只有一对进入区/退出区,也就是目标从一个摄像头的退出区离开后,只能出现在另一个摄像头的进入区。

Huang 等人^[12]提出了一种概率方法用于识别高速公路上跨摄像头下的车辆,用车辆的颜色均值描述其外观,车辆在摄像头间的转移时间符合高斯分布,并且假设初始转移概率是已知的。该方法在视频图像噪声较大的情况下依然能够取得较好的跟踪效果,然而仅采用平均颜色的外观模型无法描述类似行人这样颜色比较复杂的目标,且该方法的应用场景是车辆运动方向固定的高速公路,在其他复杂场景中高斯分布假设不一定适用。另外,车辆初始转移概率也是跟踪系统之外的先验知识,需要额外求得。

Kettner 等人^[13]将非重叠摄像头跟踪看作建立目

标一致性的线性规划问题,利用贝叶斯方法构建跨摄像头时目标的运动路径。该方法中目标运动路径和在摄像头间的转移概率是人工设定且固定的,也没有利用目标位置、速度、转移时间等信息统一建模,因此当目标运动轨迹脱离预定路线时跟踪会失败。

Ellis 等人^[14]采用非监督方式学习目标在相邻摄像头间转移的时间相关性来构建摄像头网络的拓扑图。作者分两步提取时间相关信息,首先识别每个摄像头的进入和退出区,然后统计观测实例的移动轨迹来建立一个摄像头的退出区与相邻摄像头的进入区之间的一致性关系。该方法的优点是无需计算相邻摄像头间的每个实体运动轨迹的一致性关系,而是基于统计大量轨迹的方法来建立相邻摄像头进入区和退出区的一致性关系,因此是完全无监督的,特别适合实际场景中长时间稳定运行的跟踪系统。然而由于一致性不是建立在计算单个实体轨迹上,同一个系统在跟踪诸如人和车辆等速度差别很大的目标时效果不理想。

Javed 等人^[15]提出一种利用摄像头间的时空线索和目标外观线索的贝叶斯框架实现跨摄像头跟踪。摄像头间的时空模型由 Parzen 窗从训练数据中得到,而目标经过相邻摄像头时外观的变化由颜色模型之间的距离度量。相邻摄像头的两个观测实例是同一个目标的概率就由时空信息和目标外观决定,称之为一致性概率(Correspondence Probability),跟踪过程就是对它进行最大后验概率估计。该方法不需要校准摄像头,也不需要摄像头网络的拓扑结构,因此具有较好的扩展能力。其不足之处是一致性概率的求解本身就比较复杂,而且训练数据需要手工标记运动轨迹等信息。

Xiaotang Chen 等人^[16]在 Javed 等人工作的基础上,提出了另一种非监督方法构建摄像头间转移模型,该模型同样由摄像头间的时空线索和目标外观线索两个部分组成。不同之处在于作者将时空模型描述为一个摄像头连接(Link)中进入观测实例和退出观测实例的互相关函数。互相关函数虽然比较容易实现,但它存在峰值不明显以及没有考虑连接的转移时间分布的问题。基于上述互相关函数,作者提出通过减小正相关实例的转移时间方差,以消减大量负相关实例带来的影响,从而突出互相关函数的峰值。通过不断迭代,该方法能够估计任意有效连接的转移时间分布,而且避免了求解非重叠摄像头中目标一致性问题。目标外观线

索就是对目标外观特征的描述。文中作者采用基于颜色的特征,提出利用颜色特征转移算法(CCT)^[17]来建立摄像头间外观转换模型。该算法克服了摄像头间光照变化对目标外观颜色的不利影响,缺点是不同的颜色转移方向(例如从摄像头 x 到 y 与从 y 到 x)会极大地影响 CCT 的性能。

3.2 重叠摄像头跟踪方法及性能分析

多个摄像头重叠时的目标跟踪在解决单摄像头视域中目标被遮挡、目标在视域中过小等方面有重要意义。它可以分为两个过程,首先在多个摄像头中检测到目标,然后再进行单摄像头下的目标跟踪。本节只介绍多个摄像头下目标检测的方法。

最简单的检测方法是特征匹配法^[18-19],利用目标的颜色、纹理、HOG 等特征或特征组合作为模板,对各个摄像头的视频帧进行匹配。另外,摄像头校准信息对于目标检测也有帮助^[20]。

Khan 和 Mubarak 在文献[21]中将跟踪同一个目标的一致性看作为对目标的连续标记(Consistent Labeling)。文中先自动计算具有重叠区域的每个摄像头的视域线(FOV Lines),一种在其他摄像头中观察到的该摄像头的视域边界,然后可由求得各摄像头视域间的单应性关系,从而建立目标在相邻摄像头之间的一致性。该方法不需要特征匹配和校准摄像头,而且可以知道目标在一个摄像头视域的不同区域对相邻摄像头的可见性。该方法的缺点是要求相邻摄像头的视域线是互相可见的,即使只有部分不可见也会导致一致性的置信度很低。

高翔等人^[22]提出一种基于人体中线的目标检测算法,用于重叠摄像头行人跟踪。该算法假设人体在中线两侧近似对称,通过最小中位数均方差检测各个摄像头采集的图像中候选目标的中线,然后以不同摄像头视角间的单应性关系作为几何约束定义一个中线匹配似然函数,实现在不同视角中检测目标。由于该算法需要依赖背景差分法提取前景目标,以及人体中线假设过于简单化,因此不适合背景复杂和行人姿态复杂的应用环境。

熊俊等人^[23]在研究智能停车场管理系统中,计算当前视域中候选目标的世界坐标与前一摄像头视域中目标的世界坐标之间的距离,距离最小的候选目标就是前一视域中正在跟踪的目标。世界坐标实际上就是目

标在停车场的中物理位置,因此需要事先对摄像头进行校准,在线跟踪时也需要对车辆进行精确定位。

Zhang Z 等人^[24]在研究头盔跟踪系统中,将三个布置在同一个头盔上的摄像头采集的视频图像通过几何变换缝接(Stitch)成一个广域视频,然后在这个视频中跟踪目标。由于头盔上的摄像头位置相对固定,因此很容易对摄像头进行标定,摄像头之间的单应性关系也容易计算,在此基础上缝接多个视频图像的速度也较快。该方法等同于单摄像头下的目标跟踪,省去了在不同摄像头中检测目标的复杂过程。然而,该方法只适用于多个摄像头之间间距非常小,几乎不存在视域盲区的应用环境,否则缝接好的视频图像容易形成空白区。

4 跨摄像头跟踪的发展趋势展望

与单摄像头目标跟踪相比,跨摄像头目标跟踪是一个较新的研究领域。要实现非重叠或重叠摄像头下的有效跟踪,合适的摄像头关联模型或图像融合算法是关键。在单摄像头跟踪研究中,类似 VOT 等测试和比较跟踪算法优异性的基准数据集^[25]起着重要作用,

虽然在跨摄像头跟踪领域也有一些测试数据集^[26],但大部分是关于重叠摄像头跟踪的,而更符合实际跟踪需要的非重叠摄像头跟踪的数据集非常少。开发既能在重叠摄像头、又能在非重叠摄像头下测试和比较摄像头关联模型和图像融合算法的基准数据集是迫切需要解决的问题。另外,当前大多数跨摄像头跟踪算法要么偏向非重叠摄像头下的跟踪,要么偏向重叠摄像头下的跟踪,在未来的研究中应该将重叠与非重叠摄像头跟踪统一到同一个算法框架下。

5 结语

本文先介绍跨摄像头目标跟踪的基本流程,在此基础上指出跨摄像头目标跟踪研究面临的两大挑战:建立摄像头关联模型和设计图像融合算法。然后对各类非重叠摄像头跟踪时的摄像头关联模型,以及重叠摄像头跟踪时的图像融合算法进行了综述,分析了不同方法的性能。本文最后对跨摄像头目标跟踪的发展趋势进行了展望,相信今后会出现鲁棒性更好、扩展性更强、智能化更高的跨摄像头目标跟踪算法。

参考文献:

- [1]曾巧玲,文贡坚.运动目标跟踪综述[J].重庆理工大学学报自然科学版,2016,30(7):103-111.
- [2]A Yilmaz, O Javed, M Shah. Object Tracking: A Survey. ACM Computing Surveys, 2006, 38(4):1-29.
- [3]管皓,薛向阳,安志勇.深度学习在视频目标跟踪中的应用进展与展望[J].Acta Automatica Sinica, 2016, 42(6):834-847.
- [4]Kalal Z, Mikolajczyk K, Matas J. Tracking-Learning-Detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(7):1409-22.
- [5]王津,张学杰.多摄像头视频监控系统中基于 SIFT 特征的目标匹配研究[J].云南大学学报自然科学版,2009(s1):37-42.
- [6]HUANG Y F, Chu-Yang L I, Cai-Rong Y A N. Object Tracking for Multiple Non-overlapping Cameras Based on TLD Framework[J]. DEStech Transactions on Engineering and Technology Research.2016
- [7]Castaneda J N, Jelaca V, Frías A, et al. Non-Overlapping Multi-camera Detection and Tracking of Vehicles in Tunnel Surveillance[C]. International Conference on Digital Image Computing: Techniques and Applications, 2011:591-596.
- [8]李志华,陈耀武.基于多摄像头的目标连续跟踪[J].电子测量与仪器学报,2009,23(2):46-51.
- [9]陈伟宏,肖德贵.一种非重叠多摄像头的实时监控系統[J].计算机工程与应用,2006,42(18):218-220.
- [10]Javed O, Shafique K, Shah M. Appearance Modeling for Tracking in Multiple Non-Overlapping Cameras[J], 2005, 2:26-33.
- [11]Chu C T. Robust and Consistent Human Tracking Within Camera and, Across Multiple Cameras[J], 2013.
- [12]Huang T, Russell S. Object Identification in a Bayesian Context[C]. Fifteenth International Joint Conference on Artificial Intelligence. Morgan Kaufmann Publishers Inc, 1997:1276-1282.
- [13]Kettner V, Zabih R. Bayesian multi-Camera Surveillance[J], 1999, 2(2):2253.
- [14]Ellis T J, Makris D, Black J. Learning a Multi-Camera Topology. Joint IEEE Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS), 2003:165-171.
- [15]Javed O, Rasheed Z, Shafique K, et al. Tracking Across Multiple Cameras with Disjoint Views[C]. IEEE International Conference on Computer Vision, 2003. Proceedings. 2003:952-957 vol.2.

- [16]Chen X, Huang K, Tan T. Object Tracking Across Non-Overlapping Views by Learning Inter-Camera Transfer Models[J]. Pattern Recognition, 2014, 47(3):1126-1137.
- [17]Reinhard E, Ashikhmin M, Gooch B, et al. Color Transfer between Images[J]. IEEE Computer Graphics & Applications, 2001, 21(5): 34-41.
- [18]Utsumi A. Multiple-Camera-Based Human Tracking Using Non-Synchronous Observations[J]. Proc of Accv, 2000:1034-1039.
- [19]Chang T H, Gong S. Tracking Multiple People with a Multi-Camera System[C]. IEEE Workshop on Multi-Object Tracking. IEEE Computer Society, 2001:0019.
- [20]Cai Q, Aggarwal J K. Tracking Human Motion in Structured Environments Using a Distributed-Camera System[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1999, 21(11):1241-1247.
- [21]Khan S, Shah M. Consistent Labeling of Tracked Objects in Multiple Cameras with Overlapping Fields of View[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2003, 25(10):1355-1360.
- [22]高翔, 朱婷婷, 刘洋. 多摄像头系统的目标检测与跟踪方法研究[J]. 计算机技术与发展, 2015(7):221-225.
- [23]熊俊, 陈临强. 基于多摄像头协同模式的智能停车场管理系统[J]. 机电工程, 2010, 27(4):76-78.
- [24]Zhang Z, Cheng J. Multi-Camera Tracking Helmet System[J]. 2013, 1(2):76-79.
- [25]Wu Y, Lim J, Yang M H. Online Object Tracking: A Benchmark[C]. Computer Vision and Pattern Recognition. IEEE, 2013:2411-2418.
- [26]Ristani E, Solera F, Zou R, et al. Performance Measures and a Data Set for Multi-Target, Multi-Camera Tracking[C]. European Computer on Computer Vision Workshops. 2016.

作者简介:

王先彬(1987-),男,四川泸州人,硕士研究生,研究方向为机器智能

严明(1988-),男,四川成都人,博士研究生,研究方向为机器智能

收稿日期:2016-12-29

修稿日期:2017-03-10

Survey on Multi-Camera Object Tracking

WANG Xian-bin, YAN Ming

(College of Computer Science, Sichuan University, Chengdu 610065)

Abstract:

Multi-camera tracking of objects, whose motions are separated both temporally and spatially when they move between cameras, is valuable in intelligent transportation, public security, etc. Aiming at the complexity of multi-camera object tracking, firstly presents a general framework for the tracking system, then, analyzes the main challenges on multi-camera object tracking along with its current research status in detail, and highlights camera correspondence models and image fusion algorithms are. Finally, discusses the trends in multi-camera object tracking.

Keywords:

Multi-Camera; Object Tracking; Camera Correspondence Model; Image Fusion