

多摄像机目标跟踪系统综述

陈炳文¹, 王文伟¹, 秦前清²

(1. 武汉大学 电子信息学院, 湖北 武汉 430079;

2. 武汉大学 测绘遥感信息工程国家重点实验室, 湖北 武汉 430079)

摘要: 介绍国内外对多摄像机的研究现状; 描述了多摄像机目标跟踪系统的各种构架并进行分析比较; 详细阐述各功能模块, 分别就算法研究现状、存在的问题及发展趋势进行了分析和总结; 最后对研究前景进行展望。

关键词: 多摄像机; 检测跟踪; 数据融合

中图分类号: TP391.41

文献标识码: A

文章编号: 1674-7720(2010)17-0001-03

Overview of target tracking system based on multi-camera

CHEN Bing Wen¹, WANG Wen Wei¹, QIN Qian Qing²

(1. College of Electronic Information, Wuhan University, Wuhan 430079, China;

2. State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China)

Abstract: Firstly, the developing status at home and abroad is presented. Secondly, several frameworks are described and compared. Then each functional module is analyzed and summarized with respect to the existing algorithms, the problems and the developing trend. Finally, some research issues of future study in this area are pointed out.

Key words: multi-camera; detection and track; data fusion

视频监控是计算机视觉领域的热点研究对象之一。它采用图像处理、模式识别和计算机视觉等技术, 抽取并分析视频源中的关键信息, 及时发现并处理监控场景下的异常情况。视频监控系统在民用和军事领域中都有极大的应用前景。

由于单摄像机的视野域有限, 因此在实际的目标跟踪系统中, 更多采用多摄像机系统。多摄像机的使用有利于解决遮挡、场景混乱、环境光照突变情况下的运动目标跟踪等问题, 但它同时也带来了一些新难题, 包括多摄像机之间的目标匹配、摄像机协作、摄像机之间的自动切换和数据融合等。

1 多摄像机跟踪系统构架

1.1 集中式的系统构架

在集中式系统结构^[1]中, 各个摄像机采集的视频数据直接被送到融合中心, 在那里进行数据对准、数据相关、航迹记录、预测与综合跟踪等, 而相机无自主处理能力。这种结构的特点是信息损失小、精度高, 但其对系统

通信要求较高, 融合中心计算负荷重, 系统效率低, 实时性差。因此该结构的实用性差, 生存能力低。

1.2 基于摄像机的系统构架

基于摄像机的系统构架^[2-4]主要分为传感器处理单元 SPU (Sensor Processing Unit)、中央处理单元 CPU (Central Processing Unit) 和用户接口 GUI (Graphics User Interface) 三部分。系统构架如图 1 所示。

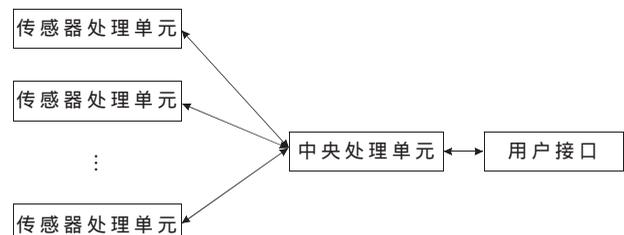


图 1 基于摄像机的系统构架

(1) 传感器处理单元。由单个摄像机和处理机组成, 摄像机可以是红外的或全方位的摄像机。该模块有自主

综述与评论 Review and Comment

处理能力,能够自动获取视频数据,并进行单摄像机内目标的检测、分类、跟踪等。然后把检测结果数据,如目标类型、位置、速度、时间戳、摄像机参数(位移、旋转、放大倍数等)等传递给CPU。

(2)中央处理单元。主要完成SPU间的信息融合,建立SPU间的通信,并进行相关信息的数据库操作。摄像机的分配调度是CPU的关键功能,根据任务的优先级、SPU的负担、摄像机的可视度因素进行分配。

(3)用户接口。用户通过GUI可以获取目标在三维场景和目前地图中的信息,也可以对检测跟踪过程进行一定的约束,如通过设置感兴趣区域(ROI)对特定的区域进行严密监控。

1.3 面向对象的系统构架

参考文献[5]提出了一种面向对象的多摄像机结构。系统主要分为检测代理SDA(Specialized Detection Agency)、处理组单元PRC(Processing Cluster)、群组管理单元CM(Cluster Manager)和用户接口GUI等。系统构架如图2所示。

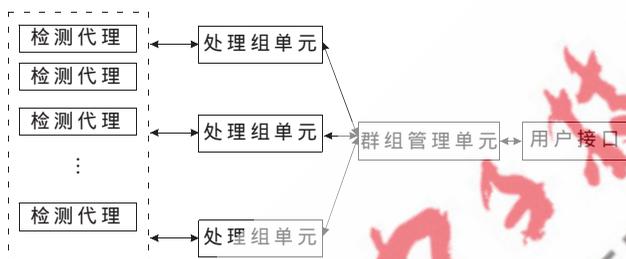


图2 面向对象的系统构架

(1)检测代理。该模块与基于摄像机结构的传感器处理单元一样,也是完成单摄像机下的处理工作,但是增加了与PRC间的接口。

(2)处理组单元。系统为每个跟踪目标分配一个处理组单元。每个PRC根据指定目标对象的情况动态控制多个SDA。动态摄像机管理单元(DSM)根据目标、SDA的参数和所估计的观测质量动态挑选SDA。全部PRC组成SDA和CM间的动态层,是最复杂的处理单元层。PRC的内部结构如图3所示。

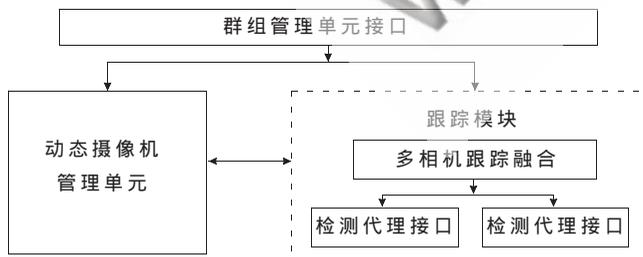


图3 处理组单元的内部结构

(3)群组管理单元。负责管理动态层的PRC,为每个目标分配最优的PRC。用户可通过CM设置PRC的参数。

(4)用户接口。该模块的功能与基于摄像机结构的

GUI一样,不再重复。

基于摄像机的结构与集中式结构的区别在于:前者的每个摄像机单元有自主处理能力,送往融合中心的数据是经过加工的精炼数据。相对于集中式系统,此类系统的可靠性高,可以减小通信量,减小带宽需求,有利于大范围的目标跟踪作业。面向对象的系统构架与基于摄像机的构架的不同之处在于它为每个目标对象单独设置一个处理单元,将多目标处理化为多个单目标处理,可扩展性强,通信量也较小。集中式的系统构架因其效率低、实时性差、生存能力低,实时跟踪系统一般不予采用。目前多摄像机目标跟踪系统架构应用较多的是分布式结构。

2 多摄像机跟踪系统的功能模块

2.1 目标检测

视频运动目标检测就是从视频序列中检测出运动区域,并对运动区域进行分类,找到感兴趣目标,如人、车等。

2.1.1 研究现状

目前提出的运动检测算法比较多,根据不同的用途和不同的环境,每种算法都有其优缺点。参考国内外文献,本文将检测算法分为四大类。

(1)基于运动场的检测。其基本思想是:用光流矢量场^[6-7]估计出每帧的运动场,然后根据每点的运动矢量分割出运动区域。该类方法可得到较好的目标边缘,局部性能好。但该方法要求目标的帧间运动量不能太大,需要很高的采样率,而实际的采集系统并不能满足此要求,且该算法复杂,计算量大,很难满足实时性要求。

(2)基于变化的检测。其基本思想是:对视频帧差图像进行检测,帧差可以是相邻帧的帧差或几帧间的帧差。该方法具有较强的场景变化适应能力,抗光照变化和抗噪声能力强,但容易产生空洞现象,目标不完整。VSAM项目提出了一种自适应背景减除与三帧差分相结合的混合算法,能解决空洞现象。

(3)基于概率的检测。其基本思想是:为像素建立概率模型,并可根据场景设定概率模型的类型和数目,从概率论的角度进行检测。该方法的理论基础扎实,可以加入先验知识,检测效果好。混合高斯法、非参数法、隐马尔科夫模型法等是其典型代表。

(4)基于模板的检测。其基本思想是:预先建立对象模板,采用模板匹配的思想实现目标检测。该类方法由于加入了对象的先验知识,对于复杂的对象检测效果更好。对象模板可以通过学习的方法建立,也可以由人机交互产生。该类方法需预先对目标进行建模,对象模板的优劣会直接影响到检测结果,所以该类算法一般适用于特定对象的检测。

2.1.2 存在的问题及发展趋势

视频检测的困难性体现在以下方面:抗光照变化

综述与评论 Review and Comment

(快或慢)、抗抖动、抗背景扰动(树木摆动、背景物体消失或出现)、抗阴影、抗颜色相似等。虽然人们研究了各种各样的检测方法来解决这些问题,并在某个方面取得了一定的效果,但目前还没有一种通用的算法。

有人结合了概率模型和预测器如卡尔曼滤波器^[12]来检测目标;有的人融合了时域、空域(特别是梯度域),如封春升提出时域和梯度域相结合的视频对象提取算法^[10],该方法结合了背景差分法、帧差分法和梯度域。

视频检测的研究重点仍是以下几个方面:优良(如自适应)的模型或模板、高效的算法、较好的预处理和后处理等。就检测的发展来看,研究较多的还是变化检测法和概率模型法,一些预测技术、时空融合技术的结合也是研究热点。

2.2 目标跟踪

目标跟踪就是对检测出来的感兴趣目标进行持续的尾随观测,获得目标的状态参数如位置和速度等,以便进行下一步的处理分析,如行为分析等。

2.2.1 研究现状

从20世纪80年代到现在,出现了众多的跟踪算法,文献^[11]将视频跟踪算法分为四类,分别是基于区域的跟踪、基于特征点的跟踪、基于变形模板的跟踪和基于模型的跟踪。本文考虑多摄像机的条件,参考相关文献将实用的跟踪算法主要分为两种:基于特征的跟踪和基于模型的跟踪。

(1)基于特征的跟踪。其基本思想是:提取目标的特征,使用匹配算法、代价准则进行匹配跟踪。针对多相机条件下,同一目标在不同视野内的观测值相差较大,故提取的特征应与视点无关。如VSAM项目组采用物体的3D轨迹、归一化色度直方图作为特征。

该类方法由于采用不变量性质的特征,故可以将特征信息传递给后续相机以实现持续跟踪,而不用考虑摄像机间的视野交叉与否,因此更符合实际情况。该类方法的核心是特征的提取,而大部分工作是单视野内的目标跟踪,故可以在单视野内使用其他更有效的算法(如区域相关法、Snake模板法等)来提高精度。

(2)基于模型的跟踪。其基本思想是:利用多相机的交叉视野建立3D模型(汽车、人)来进行跟踪。文献^[12]结合多源数据进行3D定位,并利用3D卡尔曼滤波器进行预测跟踪。该方法先利用先验知识建立目标的结构模型,再根据实际观测值得到模型参数值。

该类方法可得到精确的3D轨迹,跟踪可靠性高,但其模型的建立较难,特别是像人这样的非刚性物体,且因在3D空间上进行跟踪,运算量也较大。

2.2.2 存在的问题及发展趋势

上述的两类方法都或多或少存在缺点。显然,特征的有效提取是基于特征方法的一大难点,且该类方法在单视野内有时还要解决遮挡问题。相对而言,基于模型的方法一般无遮挡问题,但很难建立一个通用的模板

(如变形模板)。另外如何定义匹配的度量来使跟踪更精确又是一大难题。

无论哪种方法,鲁棒性、准确性、快速性都是当前跟踪技术的努力方向^[11]。融合两类方法,在建立模型的基础上提取其不变量是跟踪算法的发展趋势。

2.3 数据的融合

数据融合就是根据给定的融合算法对各个相机给出的信息完成数据配准,做出决策并进行状态更新。数据融合可分成三个部分:数据配准、数据关联和决策处理。数据配准就是将不同时间、不同视角、不同设备获得的数据变换到同一个参考框架中,使之具有可比性。大多数的数据配准方法都由四个步骤组成:特征点的提取、特征匹配、变换模型估计和数据转换。数据关联就是将配准后的信息和目标建立对应关系。决策处理就是针对每个目标的信息更新状态,调整调度策略,并给出下一阶段预测信息。

2.3.1 研究现状

由于目标跟踪中处理的数据一般是图像,参考文献^[1],本文把数据融合分为3类:像素级融合、特征级融合和决策级融合。

(1)像素级融合。该方法融合各个相机的信息,再从中提取特征进行判断识别,属于较低层的数据处理。该方法的优点是信息量损失最小,决策可信度高,但该方法计算复杂度高,抗干扰性差,不灵活。文献^[15]结合多源数据进行3D定位、预测处理的方法就是属于该类方法。

(2)特征级融合。该类方法是由每个相机自己抽取特征信息,融合中心再进行特征分析处理。该方法的数据量有了一定的压缩,有利于实时处理,其性能处于像素级和决策级之间。

(3)决策级融合。该类方法先由各相机自主做出决策,然后在融合中心完成决策的融合。该类方法的抗干扰性强,灵活性好,但信息损失量最大、精度最低,一般不宜采用。

2.3.2 存在的问题及发展趋势

上述三类方法各有优缺点,像素级融合要解决数据的标定问题,特征级融合要解决特征的有效提取,而决策级融合要提高精度就要提高各个相机决策的可信度。目前用的最多的是特征级融合和像素级融合,而决策级融合因信息损失太大而很少被采用。

本文详细介绍了基于多摄像机的目标跟踪系统,对国内外的研究现状进行了总结。描述了多摄像机目标跟踪系统的系统构架并进行了对比;详细阐述了各模块的功能,分别就算法研究现状、存在的问题及发展趋势进行了分析和总结。

随着硬件技术的逐渐成熟,多摄像机目标跟踪将是监控业未来的重要技术应用,不同的研究方向也有相应的研究重点:(1)研究多源数据融合。采用光学传感器、

红外传感器等同时进行数据的采集,或使用多分辨率的数据进行跟踪。这有助于扩展时空的覆盖范围,提高系统的鲁棒性。(2)研究客观的评价标准。系统的性能好坏需要用通用的标准进行评定,就目前而言,较多的还是使用主观的评价标准,客观的评价标准还有待研究。

参考文献

- [1] 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.
- [2] ELLIS T. Multi-camera video surveillance [C]. Security Technology, 2002. Proceedings. 36th Annual 2002 International Carnahan Conference on 2002:228-233.
- [3] Nakazawa atsushi, Kato hirokazu, Hiura shinsaku, et al. Tracking multiple people using distributed vision systems[C]. IEEE international conference on robotics and automation. Washington. DC, IEEE,2002:2974-2981.
- [4] CHANG TH, GONG S. Tracking multiple people with a multi-camera system [C]. IEEE Workshop on Multi-Object Tracking, 2001:19-26.
- [5] MONARI E, VOTH S, KROSCHEL K. An object-and task-oriented architecture for automated video surveillance [C]. AVSS'08. IEEE Fifth International Conference on 2008:339-346.
- [6] LUCAS B D, KANADE T. An iterative image-registration technique with an application to stereo vision [C]. Proceedings of the International Joint Conference on Artificial In-

telligence. Vancouver, BC, Canada: William Kaufmann, 1981,674-679.

- [7] Horn Berthold K P, Schunck Brain G. Determining optical flow[J]. Artificial Intelligence, 1981,17(123):185-203.
- [8] 钱渊,张晓燕,夏靖波.视频对象分割技术综述[J].探测与控制学报,2008,30(2):64-67.
- [9] WREN C, AZABAYEJANI A, DARREL T, et al. Pfunder: Real-time tracking of the human body[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997,19(7):780-785.
- [10] 封春升,郝爱民,何兵,等.一种时域和梯度域相结合的视频对象提取算法[J].中国图象图形学报,2008,13(3):494-498.
- [11] 侯志强,韩崇昭.视觉跟踪技术综述[J].自动化学报,2006,32(4):603-616.
- [12] JIN H, QIAN G, RAJKO S. Real-time multi-view 3D object tracking in cluttered scenes [M]. Springer Berlin/Heidelberg, 2006:647-656.

(收稿日期:2010-01-08)

作者简介:

陈炳文,男,1986年生,博士研究生,主要研究方向:图像处理、视频检测跟踪、模式识别。

王文伟,男,1966年生,副教授,主要研究方向:印刷体及手写体字符识别、图像处理及模式识别、神经网络。

秦前清,男,1961年生,教授,主要研究方向:图像处理及模式识别、空间信息压缩与渐进传输、遥感影像多尺度解译与分类识别。