

行人跟踪技术综述

马晓婷

(四川大学 计算机学院(软件)学院,四川 成都 610064)

摘要:行人跟踪是当前机器视觉中非刚性运动目标跟踪领域的热点问题,将这一问题分为行人检测和行人跟踪两大部分,并对其进行了详细介绍。分析了不同的检测和跟踪方法,对行人跟踪中存在的疑难问题进行了总结,最后对行人跟踪问题的研究进行了简单的展望。

关键词:计算机视觉;行人检测;行人跟踪;图像序列;视频监控

中图分类号:TP391 文献标识码:A 文章编号:1009-3044(2011)19-4647-03

Pedestrian Tracking Technology Overview

MA Xiao-ting

(Computer(Software) College Sichuan University, Sichuan 610064, China)

Abstract: Pedestrian tracking is a hot issue of the non-rigid motion target tracking in the current machine vision field, Divides this problem into pedestrian detection and pedestrian tracking, and two parts of this is introduced in detail. This paper analyses different detection and tracking methods. Make a summary of difficult problems exists in pedestrian tracking. Finally, the research on pedestrians tracking problem is prospected briefly.

Key words: computer vision; pedestrian detection; pedestrian tracking; image sequence; video monitoring

它作为计算机视觉领域的前沿科学,涉及到图像处理、模式识别和人工智能等学科知识,是通过对图像序列或视频监控中的行人目标进行检测、提取、识别和跟踪,获得行人的位置、速度、加速度以及运动轨迹等参数,是实现对行人行为分析及获得更深层次行为理解的重要步骤。有关行人跟踪的问题及分析已经取得了一定的进展^[1-2],由于行人的非刚性及其运动的复杂性,以及行人运动中的被遮挡和自遮挡等现象,使行人跟踪问题更加具有挑战性。

1 行人检测

行人检测属于运动目标检测,目的是从序列图像中将变化区域(行人)从背景图像中提取出来^[3]。行人检测是行人跟踪的第一步,是对行人行为进行分析和理解的基础,它在各种公共场合(如道路交通领域、居民小区、银行、停车场等场所)的应用正逐渐被推广。近几年,行人检测已经得到国内外学者的关注和研究。一般根据摄像机是否运动^[4],可以将行人检测分为两类:静态背景下的行人检测和动态背景下的行人检测^[5]。静态背景即摄像机是静止的,此时的检测只有行人相对于摄像机的运动;动态背景时是指背景变化,是摄像机和行人都存在运动,此时是摄像机和行人的复杂的相对运动。

静态背景下行人检测常用的方法有:背景差法,帧差法和光流法。

1)背景差法是摄像机静止情况下,广泛使用的检测运动目标的方法^[6]。首先获得一个背景图像,然后将当前图像与背景图像做差,根据像素差值与阈值大小判断此像素属于运动目标或背景图像。人工获取背景图像在应用中不易实现,因此大多通过背景建模获得近似的背景图像。目前背景建模算法很多,常用的有帧间差分、非线性中值滤波方法、线性卡尔曼滤波法、混合高斯模型等。背景模型的建立与更新等会影响跟踪结果。文献[7]将第一帧图像作原始背景,从第2帧开始对当前帧间差分图像进行二值化处理找到运动区域和非运动区域;然后用当前帧图像中的非运动区域部分对当前背景图像进行更新,通过60帧图像的迭代便可提取出清晰的背景。文献[8]用混合高斯模型得到运动人体的区域,并采用在线K均值算法近似估计混合高斯模型的参数进行参数的更新,降低了运算量,提高了运行速度。文献[9]用三维中值滤波方法建立背景模型,然后用背景差分法得到运动区域,简单易行,抗噪能力强。

2)帧间差分法是通过前后相邻的两帧图像的像素做差,并将差值图像二值化来获得运动物体位置和形状等信息的运动目标检测方法。其实现简单,计算量小,对光照变化不敏感,稳定性较好。但对噪声较敏感,可能在目标内部产生空洞,无法完整地提取运动目标。

使用背景差法和帧差法实现行人检测需要注意:

1)所选择的帧间时间间隔很重要。由于行人一般运动速度慢,应该选择较大的时间间隔,否则当行人在前后两帧中几乎完全重叠时,无法检测到行人。

2)阈值的选择相当关键。通常由于行人的服装颜色混搭或行人携带物品,单一固定的阈值难以获得可靠的结果,使用分级阈值或自动阈值分割对帧差图像进行二值化处理可以可靠实时地检测视频序列中的运动人体。文献[7]采用动态双阈值方法对背景差分图进行二值化操作。通过差分图像直方图中对应的噪声均值,将噪声分为两部分,利用直方图分别求出两部分的噪声模型,然后在

收稿日期:2011-05-15

作者简介:马晓婷(1985-),女,山东泰安人,硕士,主要研究方向为行人跟踪。

本栏目责任编辑:唐一东

■■■■■ 人工智能及识别技术 ■■■ 4647

中值滤波后的直方图上找到噪声分别与亮目标和暗目标分离的阈值,根据阈值二值化差分图,再用数学形态学对其进行处理,得到行人清晰完整的轮廓。

3)光流法主要是计算光流场,即在适当的平滑性约束条件下,根据图像序列的时空梯度估算运动场,通过分析运动场的变化对运动目标和场景进行检测与分割。

通常有基于全局光流场和特征点光流场两种方法。最经典的全局光流场计算方法是L-K法和H-S法,得到全局光流场后通过比较运动目标与背景之间的运动差异对运动目标进行光流分割,计算量大。特征点光流法通过特征匹配求特征点处的流速,计算量小、快速灵活,稀疏的光流场很难精确地提取行人的形状。

光流法不需要预先知道场景的任何信息,就能够检测到运动对象,而且可以处理背景运动的情况;但其抗噪能力差,而且计算复杂,很难实现实时处理。

另外,行人检测最有效的方法是直接进行分类^[10],文献[11]将行人的行走节律作为分类的一个特征。行人行走时的节律性表现为腿的摆动现象,使腿部运动具有周期性。因其属于人体自身的不变特征,所以比较有效,但是由于腿部运动的获得不一定容易和方便,因此这种分类还是有一定局限性的。但是与生物特征识别技术结合也必将成为有效的行人检测技术。

2 行人跟踪

行人跟踪就是在各帧图像中检测定位出行人。目前,常用算法有Kalman滤波算法^[12]、Meanshift算法^[13]、粒子滤波算法^[14]及动态贝叶斯网络^[15]等。Kalman滤波是基于高斯分布的线性运动状态预测方法,不能有效的处理多峰模式的分布情况;Meanshift算法和粒子滤波是基于概率统计的方法,需要迭代求解,其中Meanshift算法通常是利用目标的颜色信息,特征单一,容易发生漂移,但跟踪速度很快;粒子滤波算法存在粒子退化的严重问题,运算量通常较大。

根据跟踪方法的不同,一般将行人跟踪分为四类:基于模型的跟踪、基于区域的跟踪、基于主动轮廓的跟踪和基于特征的跟踪。

1)基于特征的跟踪主要包括特征提取和特征匹配两步。

特征提取是在原始图像中提取出最能描绘和识别行人的易用特征。提取的特征应具有代表性,特征计算应该相对简单,以及对图像平移、旋转、尺度变化等的不变性^[9]。行人跟踪中常用特征主要有颜色、高宽比、边缘、轮廓、周长、面积、质心、位置等。行人是非刚性目标,具有不规则性,对其提取的特征直接影响到跟踪的准确性。实际应用中,常选择多个特征相结合进行匹配,提高跟踪的准确性。文献[16]提出了一种基于空间边缘方向直方图的Meanshift跟踪算法,使用空间分布和纹理信息作为匹配特征,克服了传统的Meanshift算法,只利用颜色直方图作为特征容易造成跟踪丢失缺陷,实现了遮挡、和尺度缩放等复杂情况下对行人的有效跟踪。文献[17]提出使用空间位置、形状特征和颜色信息结合的方法,使用Kalman滤波预测进行行人跟踪,用一个紧密包含行人的矩形框中心表示行人的位置;使用了行人的高(L)、宽(W)、面积(A)、高宽比($R=W/L$)及紧密度($(C=4\pi A)/P$,P表示包含在行人矩形框中的像素个数)等形状特征;跟踪过程中,当形状特征不可靠时使用颜色特征,在实际场景中对单个和多个行人跟踪都具有很好的鲁棒性。

特征匹配是利用提取到的有效特征对相邻帧间的行人目标进行相似度计算,找到最相似的行人的过程。常见的相似性度量方法有欧式距离、街区距离、棋盘距离、加权距离、巴特查里亚系数、Hausdorff距离等^[5]。

2)基于区域的跟踪

算法基本思想是,先得到包含目标的模板(该模板可通过图像分割获得或预先人为确定,模板通常为略大于目标的矩形,也可不规则形状),然后在序列图像中,运用相关算法跟踪目标;也有学者根据人的生理组成将行人看做头、躯干、四肢等小区域的组合,通过跟踪各个小区域达到跟踪整个人的目标。

文献[9]提出了一种针对红外视频监控问题,基于图像运动区域的多行人跟踪算法,用三维中值滤波方法进行背景建模,使用背景差分法得到运动区域,然后根据运动区域之间的匹配准则,建立帧间运动区域匹配矩阵,实现行人跟踪。由于运动区域的检测效果不是很好,导致跟踪效果并不理想,对于行人部分遮挡时能跟踪到,完全遮挡时的处理不妥,但是文献所提出的算法对多行人跟踪的鲁棒性还是值得肯定的。

3)基于主动轮廓的跟踪

主动轮廓模型,由Kass^[18]等人于1987年提出,通常又称Snake模型。如果已知物体的先验形状信息,Snake模型发展为更为一般化的技术——变形模板^[19]。它是在一定优化准则条件下用参数化的几何模型去匹配图像中某一形状已知目标物体的模型,是建立在贝叶斯理论框架基础上的图像分割技术^[20]。由于运动目标的非刚性特性,及可变形模板所具有的性能和灵活性优点,使可变形模板在跟踪任务中引起了兴趣^[21]。目前为止,已有学者成功的将它用于跟踪人脸表情、嘴唇^[22]、手势^[23]及人脑细胞^[20]等非刚性物体的运动。但使用主动轮廓模型进行跟踪前需要手动初始化轮廓线。

4)基于模型的跟踪

算法基本思想首先由先验知识建立行人模型,然后通过匹配运动目标(行人)进行模型的实时更新^[9]。行人是非刚体,其自身运动灵活复杂,很难建立统一的行人模型用于实时性和鲁棒性很好的行人跟踪。所建行人模型是决定模板匹配能否成功的关键,是决定行人跟踪效果的关键因素。行人跟踪目前主要有三种形式的模型^[24]应用于人体跟踪,即线图模型,2D模型和3D模型^[4]。

文献[25]中行人跟踪用了基于粒子滤波的方法,提出了一种采用行人骨架模型作为状态空间模型,将转换图像的距离用于似然估计。状态空间由一个确定点坐标,棒的长度和棒与x轴的夹角组成,并对夹角规定了固定的范围。用6棒骨架表示的行人模型简单而又有效。实验结果显示该方法没有任何事先学习就达到了高度精确的行人跟踪。

文献[26]通过使用单个固定的摄像机获得灰度图像序列,实现了实时系统的行人跟踪。文中将行人模仿为矩形斑块并与卡尔曼动态行为模型结合使用。卡尔曼滤波用来估计行人参数。提出了三种行人模型:2-D动力学模型和2-D形状;2-D动力学模型和3-D形状;3-D动力学模型和3-D形状。行人模型基于室内和室外场景下的实验结果证明,3-D形状模型效果比2-D形状模型效果好

很多。但是在行人被部分或完全遮挡的情况下,系统的鲁棒性存在许多困难。如果算法结合先验知识,对遮挡情况下的鲁棒性会有所提高。

3 行人跟踪中的疑难问题

行人跟踪虽然已经取得了一定的研究成果,但是由于行人的非刚性特征及行人活动的灵活自主性,在具体研究中仍然存在一些难题。

一个主要困难是遮挡问题,针对遮挡问题,很多文献都提出了解决方案,其中文献[27]中提出使用匹配度的大小判定目标是否被遮挡,当判为严重遮挡时直接利用 Kalman 预测值进行跟踪,并再计算匹配度,当达到某一阈值时恢复正常算法跟踪。因此,在跟踪中如何对行人遮挡情况进行合理的判断是一个值得研究的问题;也可以研究建立更加合理的行人模型解决遮挡问题;或者使用检测辅助跟踪的结合方法可以提高跟踪准确度。另外,可以使用多摄像机缓解这一难题^[28],但这样同时会引入一些新的难题。

行人跟踪应用大部分处于户外场景下,光照对行人跟踪产生的影响也很大。文献[6]用混合高斯模型获得运动人体,通过使用卡尔曼滤波结合人体的颜色信息进行识别跟踪,在室外固定场景下,较好的解决了视频监控系统中的光照问题,计算速度较快,跟踪效率高。

所有跟踪问题中几乎都存在鲁棒性和实时性的内在冲突。具体应用中往往需根据跟踪环境合理权衡鲁棒性和实时性,选择不同的折中方案。检测和跟踪都是对数字图像和视频流进行操作,计算量大,难免会增加处理时间,因此优化跟踪算法,减少海量数据的计算,是提高跟踪实时性的必要手段。

4 展望

随着生物特征识别技术的发展,人脸识别技术的成熟,步态识别技术的进步,将生物特征识别用于行人跟踪技术将有助于更好的实现行人检测和跟踪。

随着网络的迅速普及和计算机硬件功能的强大,使用多摄像机进行行人跟踪,解决遮挡、场景混杂、光照变化等影响问题,依然是未来发展的必然趋势。

将视觉跟踪和听觉、激光、红外多传感器以及神经网络等的方法联合起来进行行人跟踪,扩展时空的覆盖范围,从多角度融合和处理有用信息,提高跟踪的鲁棒性。

另外,目前的行人跟踪系统都是针对特定场合研究和使用的,不具备通用性。研究更具通用性的行人跟踪系统也是今后学者努力的一个方向。

5 结束语

行人跟踪是当前机器视觉中非刚性运动目标跟踪领域的热点问题,该文根据问题的处理过程分步骤进行了详细介绍。将行人检测和跟踪技术中需要注意的问题及研究中存在的难题进行了总结,最后对未来发展前景做了简要的阐述。

参考文献:

- [1] Pentland A. Looking at people; Sensing for ubiquitous and wearable computing [J]. IEEE Trans Pattern Analysis and Machine Intelligence, 2000, 22(1): 107-119.
- [2] Gavrilu D. The visual analysis of human movement; A survey [J]. Computer Vision and Image Understanding. 1999, 73(1): 82-98.
- [3] 王亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述 [J]. 计算机学报, 2002, 25(3): 225-237.
- [4] 侯志强, 韩崇昭. 视觉跟踪技术综述 [J]. 自动化学报, 2006, 32(4): 603-617.
- [5] 张娟, 毛晓波, 陈铁军. 运动目标跟踪算法研究综述 [J]. 计算机应用研究, 2009, 26(12): 4407-4410.
- [6] Massimo Piccardi. Background subtraction techniques: a review [J]. 2004 IEEE International Conference on Systems, Man and Cybernetics, 0-7803-8566-7: 3099-3104.
- [7] 常好丽, 史忠科. 基于单目视觉的运动行人检测与跟踪方法 [J]. 交通运输工程学报, 2006, 6(2): 55-59.
- [8] 张雷, 刘冀伟, 王志良. 固定场景下的运动检测与运动跟踪 [J]. 模式识别, 2006(09): 287-289.
- [9] 郭永彩, 胡瑞光, 高潮. 基于图像运动区域的红外多行人跟踪 [J]. 激光与红外, 2009, 39(8): 891-894.
- [10] Pai C J, Tyan H R, Liang Y M. Pedestrian detection and tracking at crossroads [J]. Pattern Recognition, 2004, 37(5): 1025-1034.
- [11] Cristovbal C, Johann E, Thomas K. Walking pedestrian recognition [J]. IEEE Transaction on Intelligent Transportation Systems, 2000, 3(1): 155-163.
- [12] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems [D]. Transactions of the ASME- Journal of Basic Engineering, 1960, 82: 35-45.
- [13] Yizong Cheng. Mean Shift, Mode Seeking and Clustering [J]. IEEE Transactions On Pattern Analysis And Machine Intelligence, 1995, 17(8): 790-799.
- [14] Carpenter J, Clifford P. Improved particle filter for non linear problems [J]. IEEE Proc of Radar. Sonar and Navigation, 1999, 146(1): 2-7.
- [15] Pavlovi V, Rehg J, Cham T-J and Murphy K. A dynamic Bayesian network approach to figure tracking using learned dynamic models [J]. In: Proc IEEE International Conference on Computer Vision, Corfu, Greece, 1999: 94-101.
- [16] 王新红, 王晶, 田敏, 杨煜, 李志鹏. 基于空间边缘方向直方图的 Mean Shift 跟踪算法 [J]. 中国图像图形学报, 2008, 13(3): 586-592.
- [17] Juan Li, Chunfu Shao, Wangtu Xu, Hao Yue. Real Time Tracking of Moving Pedestrians [J]. International Conference on Measuring Technology and Mechatronics Automation, 2009, 978-0-7695-3583-8: 811-815.

(下转第 4653 页)

的。单从表面上看,分类越多,我们的数据传输次数就越少,然而,通过进一步的实验,我们得知,并非分类越多,数据恢复效果越好。当分类增多时,原本应该归为一类的节点被分为多类,使得同类数据相关性被破坏,造成恢复效果变差。真正使用时,我们需要在传输次数和恢复效果之间权衡,以使得该数据融合方法的实时性和可靠性达到最好。

参考文献:

- [1] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on Information Theory*, 2007, 53(12):4655–4666.
- [2] Luo Chong, Wu Feng, Sun Jun and Chen Chang Wen. Compressive data gathering for large-scale wireless sensor networks. In *Proceedings of the Annual International Conference on Mobile Computing and Networking*, pages 145–156, September. 2009.
- [3] D. L. Donoho. Compressed sensing. In *IEEE Transactions on Information Theory*, Apr. 2006.
- [4] G. Shen and A. Ortega. Transform-based distributed data gathering. In *IEEE Transactions on Signal Processing*.
- [4] Huang Yinghui and Li Guanyu. Descriptive models for Internet of things. In *ICICIP*, part 2, pages 483–486, August. 2010.
- [5] J. Chou, D. Petrovic, and K. Ramchandran. A distributed and adaptive signal processing approach to reducing energy consumption in sensor networks. In *Proc. of IEEE Infocom*, pages 1054 – 1062, Mar. 2003.
- [6] R. Cristescu, B. Beferull-Lozano, and M. Vetterli. On network correlated data gathering. In *Proc. of IEEE Infocom*, volume 4, pages 2571–2582, Mar. 2004.
- [7] K. Yuen, B. Liang, and B. Li. A distributed framework for correlated data gathering in sensor networks. In *IEEE Trans. on Vehicular Technology*, 2008, 57(1):578–593.
- [8] A. Ciancio, S. Pattem, A. Ortega, and B. Krishnamachari. Energy-efficient data representation and routing for wireless sensor networks based on a distributed wavelet compression algorithm. In *Proc. of IPSN*, pages 309–316, 2006.
- [9] Akkaya Kemal, Demirbas and Murat. The impact of data aggregation on the performance of wireless sensor networks. In *Wireless Communications and Mobile Computing*, Volume 8, pages 575–578, February. 2008.
- [10] E Candès. The restricted isometry property and its implications for compressed sensing [J]. In *Académie des sciences*, 2006, 346(1): 598–592.
- [11] Lee Sungwon and Pattem. Spatially-localized compressed sensing and routing in multi-hop sensor networks. In *GSN*, pages 11–20, July. 2009.
- [12] 康健, 左宪章. 无线传感器网络数据融合技术[J]. *计算机科学*, 2010, 37(4):31–35.
- [13] 宁刚. 基于压缩感知的信号重构算法研究[C]. *吉林大学硕士学位论文*, 2010.
- [14] 周灿梅. 基于压缩感知的信号重构算法研究[C]. *北京交通大学硕士学位论文*, 2010.
- [15] 李小波. 基于压缩感知的测量矩阵研究[C]. *北京交通大学硕士学位论文*, 2010.

(上接第 4649 页)

- [18] Kass M , Witkin A, Terzopoulous D. Snakes: active contour models[J]. In: Brady I M, Rosenfield A eds. *Proceedings of the 1st International Conference on Computer Vision*. London: IEEE Computer Society Press, 1987:259–268.
- [19] 李培华, 张田文. 主动轮廓线模型(蛇模型)综述[J]. *软件学报*, 2000, 11(6):751–757.
- [20] 刘皓挺, 姜国华, 王丽. 变形模板技术及其在多目标跟踪中的应用[J]. *系统仿真学报*. 2006. 18(4):1073–1077.
- [21] Yu Zhong, Anil K, Jain M P. Dubuisson-Jolly. Object Tracking Using Deformable Templates [J]. *IEEE Pattern Analysis and Machine Intelligence*, 2000, 22 (5):544–549.
- [22] Charles Kervrann and Fabrice Heitz. Robust Tracking of Stochastic Deformable Models In Long Image Sequences [J]. 1994 IEEE, 0–8186–6950–0:88–92.
- [23] 岑峰, 戚飞虎. 短程线主动轮廓跟踪算法的研究——在复杂背景和非刚性运动目标跟踪中的应用[J]. *计算机研究与发展*, 2003, 40(2): 283–288.
- [24] Aggarwal J K. Cai Q. Human motion analysis: A review[J]. *Computer Vision and Image Understanding*, 1999, 73(3):428–440.
- [25] Ashida J, Miyamoto R, Tsutsui H, Takao Onoye and Yukihiro Nakamura. Probabilistic Pedestrian Tracking Based On A Skeleton Model[J]. *Dept. of Communications and Computer Engineering, Dept. of Information Systems Engineering*, 2006(9):2825–2828.
- [26] Osama Masoud Nikolaos P. Papanikolopoulos. Robust Pedestrian Tracking Using A Model-based Approach [J]. *Artificial Intelligence, Robotics, and Vision Laboratory*. 1998:338–343.
- [27] 马丽, 常发亮, 郭立智. 基于颜色直方图匹配的非刚性目标跟踪算法[J]. *青岛理工大学学报*, 2005, 26(5):79–83.
- [28] Fujimura K, Hyodo Y, Kamijo S. Pedestrian Tracking across Panning Camera Network [J]. *Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems*, St. Louis, MO, USA, 2009:3–7.