

城市量化研究中视频人流统计分析^①

曹 诚¹, 卿粼波¹, 韩龙玫², 何小海¹

¹(四川大学 电子信息学院, 成都 610065)

²(成都市规划设计研究院, 成都 610081)

通讯作者: 卿粼波, E-mail: qing_lb@scu.edu.cn

摘 要: 在现代城市规划研究中, 聚焦人的信息的深度分析至关重要. 采用有效的视频分析技术处理和分析监控视频可以极大地扩充行人信息的基础数据, 对城市量化研究具有重大意义. 该研究方法通过拍摄一段时期同一街道视频进行相应的处理. 采用基于前向传播卷积神经网络模型的深度学习方式检测视频中指定监测区域的行人. 为确保行人信息的准确性, 故对检测到的行人进行跟踪处理, 同时添加跟踪目标丢失判断及相应处理. 最后量化行人数量、运动方向、滞留时间以及运动速度等信息数据, 进行相应的数据分析. 实验结果表明该研究方法能有效的实现量化行人信息数据, 为城市量化研究提供准确有效的数据支撑.

关键词: 城市量化研究; 深度学习; 目标跟踪

引用格式: 曹诚, 卿粼波, 韩龙玫, 何小海. 城市量化研究中视频人流统计分析. 计算机系统应用, 2018, 27(4): 88-93. <http://www.c-s-a.org.cn/1003-3254/6307.html>

Human Traffic Analysis Based on Video for Urban Quantitative Research

CAO Cheng¹, QING Lin-Bo¹, HAN Long-Mei², HE Xiao-Hai¹

¹(College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

²(Chengdu Institute of Planning & Design, Chengdu 610081, China)

Abstract: In the research of modern urban planning, the in-depth analysis of the information that focuses on human is crucial. The use of an effective video analysis technology to analyze and monitor video can greatly expand the basic data of pedestrian, which is of great significant to urban quantitative researches. This study deals with video that shot pedestrians in the same street for a period of time. Deep learning is used for detecting pedestrians in the specified monitoring area of the video based on the forward propagation convolution neural network model. In order to ensure the accuracy of the information for pedestrians, it tracks the detected pedestrians and determines whether the target is lost. Finally, it quantifies the number of pedestrians, the direction and speed of movement, the time of retention, etc., and carries out corresponding data analysis. The results show that the method can effectively quantify data of pedestrian information, then provide accurate and effective data support for urban quantitative studies.

Key words: urban quantitative research; deep learning; target tracking

1 引言

在中国城镇化路径转型和信息通信技术高速发展的背景下, 人本尺度的城市量化研究日益受到规划界的重视, 获取多源而维度丰富的行人信息至关重要^[1].

记录城市公共空间行人信息的天网系统视频大数据是城市规划领域潜在的数据来源, 具有信息完整、维度丰富和粒度可调三大优点, 可以对城市公共空间中高度动态化的人类活动场景进行精细化描述, 具有巨大

^① 基金项目: 成都市科技惠民项目 (2015-HM01-00293-SF); 国家自然科学基金 (61471248)

收稿时间: 2017-07-26; 修改时间: 2017-08-14; 采用时间: 2017-08-28; csa 在线出版时间: 2018-03-31

的研究前景^[2,3]。

目前城市量化研究中需要的行人数据包括经过人数、滞留时间、移动方向、年龄结构、性别、情绪、活跃程度等。其中年龄结构、性别、情绪、活跃程度等学术界研究成果相对不成熟,而人数、滞留时间、方向等则可以通过对个体进行检测、跟踪来实现,相对较准确且成熟。目前对于群体量化分析主要采用基于密度检测算法来实现大致的群体人数估计^[4],却无法分析行人个体信息。本文选取了与目前的研究热点“城市街道活力评估”、“空间品质评价”等最相关的3种数据——人流量、滞留时间以及移动方向,通过以个人为本来进行城市量化研究^[5,6]。

为了充分利用视频数据所蕴含的行人信息,本文结合基于深度学习的目标检测技术及基于核相关函数的目标跟踪技术,实现了监控区域中人流量、滞留时间、移动方向等多种参数的分析,为城市定量化研究提供准确有效的数据支撑。

2 系统综述

2.1 系统整体介绍

为了获取量化分析的行人数据,需要首先检测视频中的行人,然后为确保行人的唯一性,需要对行人进行跟踪操作。本文设计的系统如图1所示。

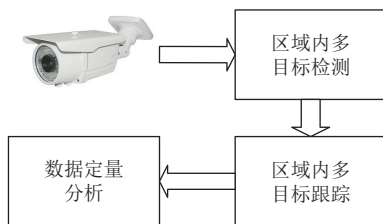


图1 系统方案示意图

指定摄像头采集视频中的监测区域,检测指定区域内的行人目标,为了能更好的达到记录行人目标的信息的效果,本系统添加对检测的行人目标进行跟踪操作,同时记录对应的量化数据信息,最终将记录的信息进行分析。

2.2 区域内多目标检测

由于只研究视频中感兴趣区域的行人信息,因此本文需要指定监测区域(即感兴趣区域),类似于图2中多边形所示。

传统方法在行人检测方面精度虽然高,但面对真

实视频的实时性、背景复杂多样、场景光照强弱变化、行人衣着和姿态多样化以及拍摄角度多样化等因素时,传统方法存在明显不足。



图2 指定监测区域

基于前向传输卷积神经网络的 Single Shot Multi-Box Detector 网络(简称 SSD 网络)^[7]在 VOC2007 测试图像中,输入 300×300 大小图像时,SSD 网络达到 72.1% mAP 的准确率,58 fps,符合实时高精度的要求。故本文采用 SSD 网络来进行行人检测。

通过预设不同长宽比的滤波器,预测边界框中的对象的类别与偏移,并将滤波器应用于网络后期的多个特征图中,用于实现多尺度检测。通过以上操作,即使对于低分辨率的图像,SSD 网络也能实现高精度与快速检测。对于本文视频中不同尺度的行人目标有很好的检测适用性。

为了能够准确的判断 SSD 网络检测的行人是否在指定的监测区域,本文采用射线法^[8]进行判断。

射线法从被判断点向无穷远处作平行于水平横轴的射线,计算该射线与多边形各边的交点个数。若交点个数为奇数则该点位于其多边形内部,若没有交点或者交点个数为偶数则该点位于其多边形外部。

2.3 区域内多目标跟踪

当判断检测的行人目标是在指定的监测区域内部时,为获取所需的信息,接下来需对检测的行人目标进行跟踪操作。

为了满足系统的需求,结合跟踪算法在跟踪速度和跟踪性能上要求,本文采用基于核函数的相关滤波器跟踪算法(Tracking Algorithm with Kernel Correlation Filter,简称 KCF 跟踪算法)^[9]。

KCF 算法总体思想为通过提取目标周围区域的图像,巧妙使用循环矩阵方式构造大量正负样本用于训练分类器,利用循环矩阵在傅里叶空间可对角化的性质将矩阵的运算转化为向量的哈达玛积(Hadamard

Product) 来对候选区域进行评估,其中响应最大的候选区域即为目标所处的新位置,最后对分类器参数重新训练以及更新。

KCF 跟踪算法跟踪速度快,跟踪效果好。然而 KCF 跟踪算法主要针对于单目标跟踪,而本文需要进行多目标跟踪,本文通过将 SSD 网络检测到的多行人目标分离,对每一个行人目标进行 KCF 算法的跟踪,实现了将 KCF 跟踪算法用于该场景下的多行人目标跟踪^[10,11]。

鉴于实际城市街道场景中背景相对比较复杂,结合所采集的视频,应用原始 KCF 跟踪算法,当行人目标相互遮挡严重时,会出现 KCF 跟踪算法跟踪目标的丢失情况。因此需要一种有效的方式来评估 KCF 跟踪算法对于目标跟踪的准确性。本文引入峰值旁瓣比 (Peak to Sidelobe Ratio, PSR) 用于评估目标 KCF 跟踪算法跟踪准确性^[12]。KCF 中循环样本找到的最优滤波器与候选区域的检测响应输出分布模型如图 3 所示,其中响应最大值即为峰值 g_{max} 。

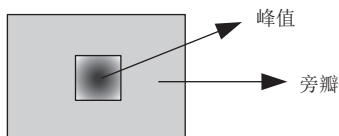


图 3 PSR 模型

PSR 值计算原理: 将响应峰值 g_{max} 中心 11×11 矩形区域之外的其余区域称为旁瓣区域,同时计算出旁瓣区域的均值 μ 和方差 σ 。利用式 (2) 即可计算出 PSR 值。

$$PSR = \frac{g_{max} - \mu}{\sigma} \quad (1)$$

设置 $PSR_{threshold}$ 阈值,当 $PSR \geq PSR_{threshold}$ 时,表示该跟踪目标是可靠的,继续跟踪该目标;反之则认为该跟踪目标发生严重遮挡,判定该跟踪目标丢失。

2.4 多目标检测与跟踪的实现

SSD 网络检测到的多行人目标如何与 KCF 跟踪算法正确结合实现是本文的重点工作。

首先指定监测区域,接下来进行 SSD 检测多行人目标,同时初始化 KCF 跟踪对象。在接下来读入的帧中,先进行 SSD 检测行人目标,然后获取当前帧 KCF 跟踪的多行人目标,判断行人目标是否还在指定的检测区域内,同时结合 PSR 判断目标是否丢失,放弃

跟踪丢失的行人目标。下一步判断 SSD 检测的多行人目标是否均与 KCF 跟踪的多行人目标相同,将不同的目标进行 KCF 对象初始化,依次执行。具体流程可参考图 4 多目标检测与跟踪的实现流程图。

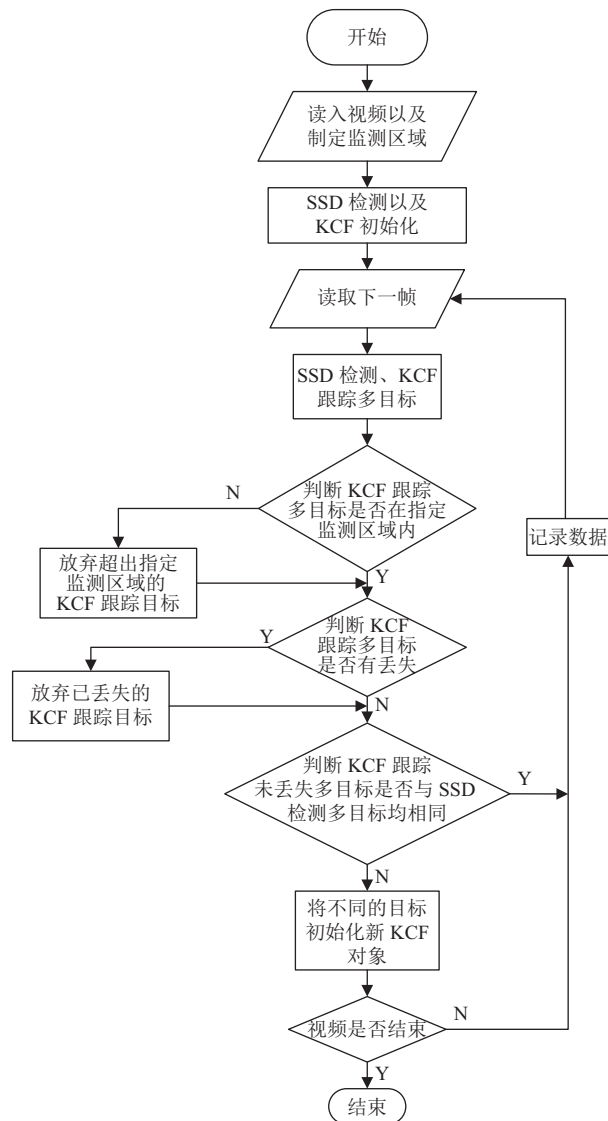


图 4 多目标检测与跟踪的实现流程图

2.5 参数计算

每次指定监测区域后,检测从视频中读取的第一帧图片,将检测到的所有行人目标坐标信息用于初始化 KCF 算法对象的分类器模型。

在读取视频下一帧后,首先记录检测当前指定区域内行人坐标信息;接下来更新 KCF 算法对象对当前帧的跟踪,同时判断目标是否仍在指定的监测区域里,计算目标在指定的监测区域里水平方向上累计的移动

距离 τ_n .

(1) 累计移动距离计算公式如式(2)(3):

$$\Delta\tau_n = \tau_{n2} - \tau_{n1} \quad (2)$$

$$\tau_n = \sum \Delta\tau_n \quad (3)$$

其中 $\Delta\tau_n$ 表示第 n 个行人目标当前帧与前一帧重心 X 轴距离差,用于实时的显示行人目标的移动方向. τ_{n2} 为行人目标当前帧指定监测区域的重心 X 轴位置, τ_{n1} 为行人目标前一帧指定监测区域的重心 X 轴位置. τ_n 为行人目标在指定监测区域内重心 X 轴位置累计改变值,用于记录行人目标在指定监测区域内总体的移动方向.

每一秒钟记录当前帧中指定监测区域里总目标数量、目标的移动方向、目标在该监测区域所滞留时间和目标在该监测区域移动速度以及当前视频帧的真实世界的时间.

(2) 滞留时间计算公式如式(4):

$$\Delta T_n = T_{n1} - T_{n2} \quad (4)$$

其中 ΔT_n 表示第 n 个行人指定监测区域的滞留时间. T_{n1} 表示该行人进入指定监测区域的时间, T_{n2} 表示该行人离开指定监测区域的时间.

(3) 速度计算公式如式(5):

$$V_n = \frac{D}{\Delta T} \quad (5)$$

V_n 表示第 n 个离开指定监测区域的行人目标经过该指定监测区域得速度值. D 表示该指定监测区域的长度.

接下来依次读取下一帧直至视频结束或者重新指定监测区域.

3 结果分析

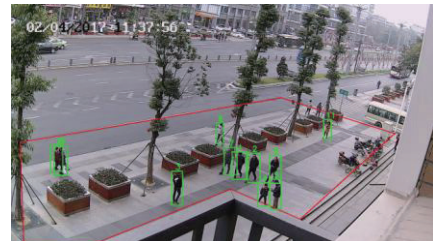
为了能够更好的为城市量化研究提供准确有效的数据支撑,本文于成都市某街道拍摄480小时的视频,对其中某一天24小时视频进行处理.其中视频分辨率1280×720,帧率20,视频画面左上角对应拍摄的世界时间信息.测试平台:惠普Z820工作站、Intel Xeon E5处理器、32G内存;Nvidia Titan X显卡、12G显存.

3.1 区域内行人检测与跟踪

本文结合街道场景,设置以下的指定监测区域,得

出如图5所示的实际效果图.

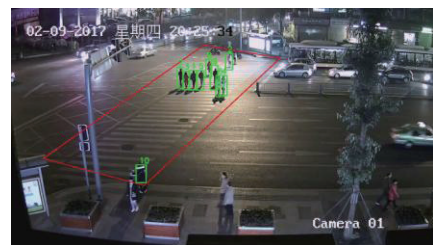
其中图5(a)中很好的检测到指定监测区域内11个行人目标,并对其进行标号1-11.从图5(b)中可以看出目标都成功跟踪,其中9号已被完全遮挡,但实现跟踪.图5中(c)(d)子图验证不同街道场景下本文方法检测与跟踪效果,可以看出,图中指定监测区域内行人目标已成功检测与跟踪.



(a) 区域1检测图



(b) 区域1跟踪图



(c) 区域2检测图



(d) 区域2跟踪图

图5 行人检测与跟踪

经过测试,平均处理一帧的时间为49ms,平均帧率20.4fps,达到实时处理效果.

3.2 行人数据分析

本文通过检测跟踪指定监测区域内的行人目标,

分析其在指定监测区域内的运动信息, 得出如图 6、图 7、图 8 所示的信息图。

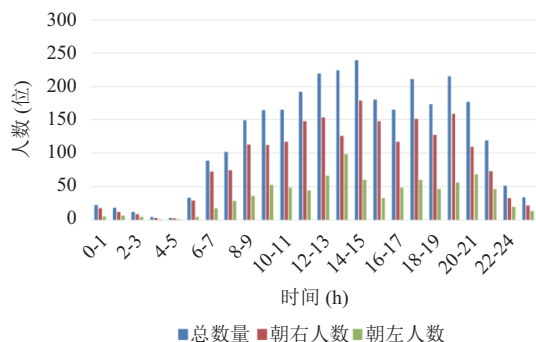


图 6 24 小时行人的总数量和移动方向

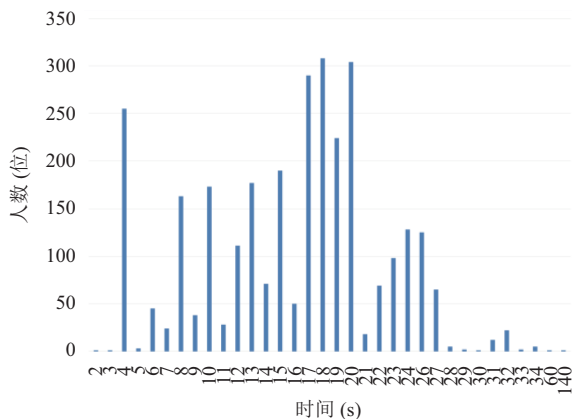


图 7 行人逗留时间分布

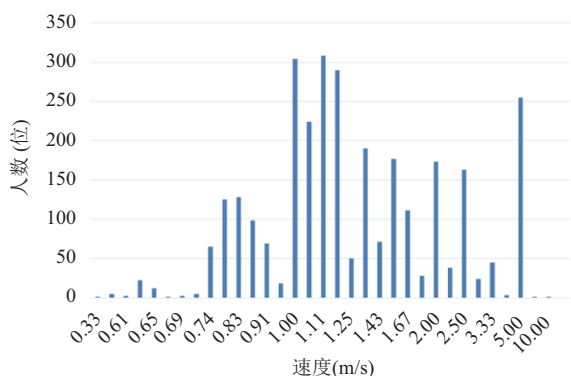


图 8 行人速度分布

图 6 中可以清楚看到 24 小时内通过该指定监测区域内行人的总数量以及移动方向, 可以获知何时通过该指定监测区域的行人数量最多, 行人移动哪个方向上数量最多. 该信息可用于城市街道活力分析中的定量分析, 以及提供了活力研究的定向数据, 为进一步深入的研究街道活力提供了较为具体的数据。

结合图 7, 得出行人通过该指定监测区域的平均时间为 16.02 s. 结合监控视频和行人通过指定监测区域时间得出: 1) 通过时间为 10 s 以下的行人基本上是从花坛边横穿过或者骑车, 其中通过时间为 4 s 的基本上都是骑车快速通过; 2) 通过时间为 10 s 至 22 s 之间的行人采用正常速度; 3) 通过时间为 23 s 以上的行人多数为散步或者一群人. 其中通过时间为 60 s 和 140 s 的为环卫工人。

结合图 7 得出图 8 所示的行人通过该指定监测区域的速度图以及实际的距离, 平均速度为 1.25 m/s.

根据实时记录行人移动轨迹方向和视频信息, 得出如表 1 所示数据。

指标	行人总数	轨迹朝右	轨迹朝左
数量	2960	2102	858

通过对视频的分析, 发现在视频右上角的位置上有个公交车站 (站牌被遮挡), 以及右侧有小区入口. 因此行人朝右的数量明显比朝左的数量要多。

4 结论与展望

本文采用深度学习, 同时结合多目标跟踪的方式, 实现对监控系统视频中城市街道中的行人信息定量和定向的分析, 为城市研究者开展城市量化研究工作提供一定意义上的参考. 然而本系统存在当出现行人遮挡时, 会出现漏检情况的问题, 有待于进一步的研究. 同时结合城市规划需求, 对数据做出进一步的分析。

参考文献

- 龙瀛. 街道城市主义新数据环境下城市研究与规划设计的新思路. 时代建筑, 2016, (2): 128-132.
- 唐婧娴, 龙瀛, 翟炜, 等. 街道空间品质的测度、变化评价与影响因素识别——基于大规模多时相街景图片的分析. 新建筑, 2016, (5): 110-115.
- Khan SD, Vizzari G, Bandini S, et al. Detecting dominant motion flows and people counting in high density crowds. Journal of WSCG, 2014, 22(1): 21-30.
- 乐勇. 基于监控视频动态信息分析的人流密度检测算法的研究[硕士学位论文]. 南京: 南京邮电大学, 2015.
- 姜蕾. 城市街道活力的定量评估与塑造策略[硕士学位论文]. 大连: 大连理工大学, 2013.
- 郝新华, 龙瀛. 街道绿化: 一个新的可步行性评价指标. 上

- 海城市规划, 2017, (1): 32–36, 49.
- 7 Liu W, Anguelov D, Erhan D, *et al.* SSD: Single shot MultiBox detector. Proceedings of the 14th European Conference on Computer Vision-ECCV. Amsterdam, The Netherlands. 2016. 21–37.
- 8 苗春葆. 点与多边形关系的射线法. 电脑编程技巧与维护, 2008, (3): 56–58.
- 9 Henriques JF, Caseiro R, Martins P, *et al.* High-speed tracking with kernelized correlation filters. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(3): 583–596. [doi: [10.1109/TPAMI.2014.2345390](https://doi.org/10.1109/TPAMI.2014.2345390)]
- 10 Ess A, Leibe B, Schindler K, *et al.* A mobile vision system for robust multi-person tracking. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, AK, USA. 2008. 1–8.
- 11 Wu HL, Li WH. Robust online multi-object tracking based on KCF trackers and reassignment. Proceedings of the 2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP). Washington, DC, USA. 2016. 124–128.
- 12 Bolme DS, Beveridge JR, Draper BA, *et al.* Visual object tracking using adaptive correlation filters. Proceedings of the 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). San Francisco, CA, USA. 2010. 2544–2550.